ISOLATION AND CHARACTERIZATION OF INTRON-MUTANTS OF THE GENE CODING FOR SMALL SUBUNIT OF RIBONUCLEOTIDE REDUCTASE OF T4 BACTERIOPHAGE

THESIS

submitted for the award of the Degree of

Doctor of Philosophy

IN

BIOCHEMISTRY

BY

ASAD ULLAH KHAN

FACULTY OF LIFE SCIENCES
ALIGARH MUSLIM UNIVERSITY
ALIGARH (INDIA)

1998

TELEPHONE : 400741
TELEX : 564—230 AMU IN

DEPARTMENT OF BIOCHEMISTRY
FACULTY OF LIFE SCIENCES
ALIGARH MUSLIM UNIVERSITY
ALIGARH—202 002 (INDIA)

*Ref. No.*_____ *Dated*_____

CERTIFICATE

This is to certify that the research work embodied in this thesis entitled "ISOLATION AND CHARACTERIZATION OF INTRON-MUTANTS OF THE GENE CODING FOR SMALL SUBUNIT OF RIBONUCLEOTIDE REDUCTASE OF T4 BACTERIOPHAGE " is an original work, unless otherwise stated, carried out by Mr. Asad Ullah Khan under my supervision and is suitable for submission for the award of Ph.D. degree in Biochemistry.

Dated: 19th August, 98

(Prof. Masood Ahmad)

DEDICATED
TO
MY MOTHER

Acknowledgements

It gives me immense pleasure to owe my humble gratefulness to my supervisor Prof. Masood Ahmad for his indispensable guidance, constructive criticism and unceasing encouragement at every step which enabled me to make my work a humble success.

I feel short of words to acknowledge Dr. Sunil K. Lal, Research Scientist, ICGEB, New Delhi, for his magnanimous generosity and unflagging interest which helped me in all my endeavours.

I am grateful to Prof. M. Saleemuddin, Ex-Chairman, and Prof. S.M.Hadi, Chairman Deptt. of Biochemistry for providing me the necessary facilities to carry out this research.

My thanks are also due to (Drs) Prof. A.N.K.Yousfi, Qayyum Husain, Jawaid Iqbal, Ms. Nahid Bano, Ms. Bilquis Bano, Fahim A. Khan, Ms. Mana Mirza, Ms.Rana Zadi, Ms.Farah Khan for the innumerable discussions and their qualitative arguments and justifications.

I am also thankful to (Drs) Saad Tayyab, Rizwan A. khan, and Mohd.Owais for their fruitful advice, amicable help and cooperation.

I also pay lot of thanks to my research colleagues (Mr.) Shams, Shakeel, Haseeb, Rizwan, Moyad, Sandeep, Salman, Yougesh, Mateen, Mushahid, Farhan, Saurabh, Athar, Anees,(Ms) Neelam, Maria, Hemlata, Neelofar, Atia for their help and cooperation during the course of Ph.D. programme.

I owe a lot to all my friends especially Dr. Abdur-Rehman, Mr. Sibghat Ullah and Mr. Khaliq.

I have no words to express the infinite gratitude to all my family members especially Mamu-mian. I also warmly acknowledge Mr. Wasif Ullah Khan who contributed significantly with their affection and moral support.

I am also thankful to the staff of Distribution Information Sub-Centre especially Mr. S. Faisal Maqbool, for searching some of the important references incorporated in bibliography.

I am thankful to University Grant Commission for the financial assistance.

ASAD ULLAH KHAN

CONTENTS

LIST OF ABBREVIATIONS

ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
CAM	Chloramphenicol
C	Cytosine
Cd	Cytidine deaminase
CR	Cytidine
cpm	Counts per minute
dAMP	Deoxy adenosine monophosphate
dCMP	Deoxy cytidine monophosphate
dGMP	Deoxy guanosine monophosphate
d_2H_2O	Double distilled water
dHTP	Hydroxymethyl cytidine nucleotide
dTMP	Deoxy thymidine monophosphate
dUMP	Deoxy uridine monophosphate
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine-tetra acetate
EMS	Ethylmethane sulphonate
EtOH	Ethanol
EX	Exon
FH2	Dihydrofolate
FH4	Tetrahydrofolate
frd	Dihydrofolate reductase
GPTG	β-Glycerophosphate tris glucose

HA	Hydroxylamine
hr	Hour
IVS	Intervening sequence
Kb	Kilobase
LB	Luria broth
M	Molar
MR	Marker rescue
$MgAc_2$	Magnesium acetate
min	Minute
µl	Microlitre
µg	Microgram
NaAc	Sodium acetate
Na-citrate	Sodium citrate
nrdB	β sub-unit of ribonucleotide reductase
nts	Nucleotides
ORF	Open reading frame
^{32}P	Radiolabled phosphorus
PAGE	Polyacrylamide gel electrophoresis
PFU	Plaque forming unit
ppk	Purified picking
rpm	Revolution per minute
RNA	Ribonuleic acid
SDS	Sodium dodecyl sulphate
T4	T4 bacteriophage
TE	Tris-EDTA
td	Thymidylate synthase
U	Uracil
UR	Uridine
wt	Wild type

LIST OF TABLES

LIST OF TABLES

LIST OF TABLES

LIST O F ILLUSTRATIONS

PREAMBLE

The *nrdB* gene of bacteriophage T4 contains 598 base-pair self-splicing intron. Several studies revealed that this intron is closely related to the group I type introns of eukaryotic organelles, such as mitochondria, chloroplast, and nucleus (Gott et al, 1986; Sjoberg et al, 1986; Cech, 1988). Autocatalysis of these introns is achieved by folding of the intron into a characteristic secondary structure necessary for intron excision and exon ligation (Belfort, 1990; Cech, 1990 Wittop-koning and Schumperli, 1994). Basic features of the predicted three dimensional tertiary structure have been characterized by mutant analysis *in vitro* and by the use of specifically positioned photochemical cross linking, affinity cleavage reagents as well as by studies with enzymatic methods, chemical probe, NMR spectroscopy and computer algorithms (Allain and Varani, 1995; Gultyaev et al, 1995).

The T4 introns are capable of self-splicing and contain the conserved structural features folded into a common secondary structure of group I type introns (Hall et al, 1987; Heuer et al, 1991). To examine the predicted secondary and tertiary structure of group I type introns, random mutagenesis utilizing novel technique with T4 are very useful. One such approach with T4 is based on a phenotype screening method (Lal and Hall, 1993). The advantage of this technique is the easy mapping of mutations due to the promiscuous recombination system of T4.

The action of group I type introns of T4 phage was found to be similar as that of ribozymes (Cech, 1987; Michel and Westhof, 1990) which have gained tremendous popularity during the recent years due to their diversified types of applications in medicine (Sarver et al, 1990; Dropulic et al,

1992; Koizumui et al, 1995).

We have therefore, selected bacteriophage T4 system which provides a convenient model for studying both the genetic and functional aspects of splicing and the major objectives of the present investigation were, (1) to study structural functional relationship of *nrdB* intron of the gene coding for small subunit of ribonucleotide reductase of bacteriophage T4 by isolating splicing defective point mutations, and (2) to study the secondary and tertiary structures of group I self-splicing intron of *nrdB* gene by isolating intra- and extra-genic second site revertants of previously characterized *frd*1 double mutants.

In the first chapter of the thesis, we have presented the literature relevent to the RNA splicing with special reference to group I introns and their special features along with the scope of the work, in view of the potential utility of ribozymes to introduce with the subject.

The second chapter describes general methodological aspects and working chemicals , bacterial and phage strains, etc. employed during the course of investigations.

The isolation and characterization of splicing defective point mutations within the *nrdB* intron of T4 phage have been described in third chapter.

Fourth chapter deals with the study of some intragenic and extragenic revertants of *frd*1*nrdB* double mutants of T4 phage.

General discussion and conclusions drawn from the experimental data have been presented in fifth chapter. The purpose of the general discussion is to coordinate very briefly the entire data and to propose a model of tertiary structure for the *nrdB* intron of Bacteriophage T4.

In the end, a brief summary of the work and bibliography are documented.

CHAPTER I: Review of literature

Genes are found to be interrupted by non-coding stretches, the so called introns in almost all classes of organisms. Therefore, processing of the genes has to be carried out in order to express them smoothly. Interrupted genes represent a minor proportion of the genome of the lower eukaryotes, but the vast majority of genes in higher eukaryotic genomes are interrupted. The primary transcript has the same organization as the gene, and is sometimes called the pre-mRNA. Removal of these intervening sequences from pre-mRNA leaves a typical messenger RNA by the process of the so called RNA splicing. In eukaryotes it occurs within the nucleus. The process of splicing together with other modifications is called post-transcriptional modification. The processed mRNA is now ready for the translation process (Sharp, 1987; Cech, 1988).

In general, prokaryotic genes do not have introns because its genome tend to be smaller and replicate faster than eukaryotic genomes. Therefore, it is thought that introns and other non-coding DNA sequences have been kept out of prokaryotes. But an extreme exception is the occurrence of three introns in bacteriophage T4 (Belfort, 1990).

DISCOVERY OF INTRONS

Introns were discovered in 1977 and heralded a new era in the study of molecular biology of eukaryotic gene expression (Berezney and Coffey, 1977). The complexity of gene organization, the combinatorial possibilities of assembling different coding exons from an RNA precursor, and the novelty of the RNA splicing process clearly indicated that eukaryotic molecular

biology would be fascinatingly different from that of prokaryotic system (Green, 1986; Padgett et al, 1986). But recent report by Morrisy and Tollervey (1995) highlights the striking similarity between prokaryotic (Doolittle,1995) and eukaryotic systems. Their evidence for a common origin of ribosomal RNA processing in eukarya, archaea and bacteria can be shown by the fact that the pre-rRNA processing has now been found in archaeon, *Solfolobus acidocaldarius* (Poter et al, 1995).

Group I introns have been discovered in mitochondrial precursor RNA genes of fungi, plants, chloroplasts and bacteriophages (Cech, 1990), and recently in tRNA gene of eubacteria (Reinhold-Hurek and Shub, 1992). Based on all known group I intron sequences , comparative folding analyses and computer search for structure with the minimum free energy of folding were performed and a generic secondary structure was derived (Wittop-Koning and Schumperli, 1994).

The conservation of short sequences among group I fungal mitochondrial introns was first noted in 1982 by Burke and Rajbhandary, Michel and coworkers as well as Davies and his associates independently. The latter two groups further showed that these introns, although quite different in primary sequences, could all be folded into a secondary structure that had a common core region. Extra nucleotides , often including long open reading frames, form stem-loop structures that are peripheral to the core. The same core structure is shared by nuclear rRNA introns in Tetrahymena, Physarum, and Pneumocystic chloroplast tRNA introns in higher plants, chloroplast mRNA and rRNA introns in Chlamydomonas, and mRNA introns in bacteriophages T4 and SPO1. Most of the conserved nucleotides exist within the four short sequences P,Q,R and S. These sequences al-

ways occur in the same order along the intron (5'-P-Q-R-S-3') but the distance between adjacent sequence elements can vary from approximately twenty nucleotides to many hundreds of nucleotides. The only other nucleotides known to be highly conserved among group I introns are the U preceding the 5' splice site and the G to which it pairs a G.C base pair in P3 and G preceding the 3' splice site (Cech, 1988).

TYPES OF INTRONS

Introns have been categorized into four major groups on the basis of their conserved nucleotide sequences and RNA structures located within and adjacent to intron (Cech, 1988). These are group I, group II , nuclear mRNA and nuclear tRNA (Fig. 1).

The group I and group II introns are particularly interesting from the view point of RNA structure because they fold to form active sites to accomplish their own splicing (Kruger et al, 1982; Peebles et al,1986 ;Van der Veen et al, 1986; Kim and Cech, 1987; Michel and Westhof, 1990).

Group I introns are further classified on the basis of distinctive structure and sequence features into four sub-groups, designated IA, IB, IC and ID (Michel and Westhof, 1990). Group IA introns for example, contain two extra pairings, P7.1/P7.1a or P7.1/P7.2 between P3 and P7, whereas many group IB and IC introns have a large extension of P5, termed P5abc (Fig. 2b). Individual introns may contain additional sequences, including open reading frames (ORF) in positions that do not disrupt the conserved core structure (Saldanha et al, 1993).

Group II introns are present in fungal, plant mitochondria and plant chloroplasts as well as in prokaryotic genomes (Michel et al, 1989a; Ferat and Michel, 1993). These are also self-splicing introns (Peebles et al, 1986; Schmelzer and Schweven, 1986; Van der Veen et al, 1986). But splicing mechanism is slightly different from that of group I introns. The self-splicing of group II introns occurs in the absence of a nucleotide cofactor since 2'hydroxyl group at the branch site participates in the first trans-esterification reaction (Peebles et al, 1986; Van der Veen et al, 1986).

The splicing process of nuclear mRNAs (Type III introns) is closely related to that of RNA-catalyzed self-splicing reaction (Sharp, 1985; Cech, 1986). Any catalytic RNA involved in the splicing of mRNA precursors could not be part of the precursor RNA, since mutational analysis has shown that the only intron sequences essential for splicing are the limited consensus sequences near the 5'and 3'splice sites (Wieringa et al, 1983). The obvious candidates for specifying this hypothetical catalytic RNA structure are the highly evolutionary conserved small nuclear RNAs (Lernar and Steitz, 1981; Maniatis and Reed, 1987). These RNAs are present in the nucleus as ribonucleoprotein particles (Lerner and Steitz, 1981; Busch et al, 1982). Type IV introns i.e. nuclear tRNAs possess the endonuclease activity and require ATP as energy source (Lambowitz, 1989).

SECONDARY STRUCTURE OF RNA

Structure - Functional Relationship: Sequence and secondary structure comparisons among the group I introns have uncovered highly conserved features, which are likely to guide RNA folding and thus be important in splicing (Davies et al, 1982; Michel et al, 1982; Cech et al, 1983).
Moreover, the cis - dominant intron mutations that disrupt splicing of yeast

mitochondrial RNA as well as second-site revertants of these mutations have implicated intron elements that are functionally important (Anziano et al, 1982; De La salle et al, 1982, Weiss-Brummer et al, 1983; Holl et al, 1985). Furthermore, a spontaneous splicing defective mutation in Tetrahymena large rRNA gene has been isolated in the heterologous *Escherichia coli* (Waring et al, 1986). Other approaches have relied on site-directed mutagenesis (Burke et al, 1986; Waring et al, 1986), or use of artificial substrates *in vitro* splicing reactions (Sullivan and Cech, 1985; Garriga et al,1986) to ascribe functional relevance to specific residues within ribozymes of Tetrahymena and Neurospora.

SECONDARY STRUCTURE OF T4 INTRONS AND THEIR RELATIONSHIP TO EUKARYOTIC INTRONS

The core structure of group I introns is based on both local and long range complementary base-pairing region (Davies et al, 1982 Cech et al, 1983; Michel and Dujon, 1983). This model of core structure of group I introns holds good in all members of this group (Waring and Davies, 1984; Michel and Cumming, 1985) . The three T4 introns bear a remarkable resemblance, both in secondary structure and in conserved sequence elements ,to the group I introns of eukaryotes. They belong to a subset, Group IA, that is characterized by systematic variations in the highly conserved sequences P,Q,R and S and by one (or, rarely two) extra stem loop structures between the P7 and P3 pairing (Michel and Dujon, 1983; Waring and Davies, 1984).

Identification of P and Q sequences (parts of which pair to form the P4 stem) as well as the R and S sequences (parts of which pair to form the P7 helix) is critical to generate the correct core structure (Fig. 2a). Although P

and Q sequences deviate from the consensus, compensatory changes allow the two sequences to pair in the expected way for each of the T4 introns (Shub et al, 1988). Now the structure for the td intron predicts P and Q sequences that are different (Shub et al, 1988) from those originally assigned by Chu et al (1986) but agree with their recent reassignments (Chu et al, 1987). The T4 introns conform to the group I intron core structure not only with respect to the central P3, P4, P7 pairing , but also with respect to the intron-exon alignment (P1 pairing). The internal guide squence (IGS), near the 5' end of the intron, has been proposed to precisely align the 5' splice site by pairing with the upstream exon. The terminal uridine of exonl always pairs with a guanosine in the IGS. These requirements are exactly met in all the three intron models. It also shows the striking similarities among the T4 introns. First, each contains two stem-loop structures (P7.1 and P7.2) between the R sequence of P7 and P3 pairing consistent with their status as group IA introns. Second, the three introns have in common two additional stem-loop structures after P9. Third, there is extensive conservation of sequences at equivalent positions of the core structure, extending well beyond the sequences that are conserved in all group I introns. Among these, conserved regions are the P7.2 stem-loop and its unpaired flanking regions, the P9 and P9.2 stem loop structures and the unpaired residues extending to the 3'end of the introns. While this homology clearly reflects the common ancestry of the T4 introns, it is interesting that these sequences have diverged less than other structural elements (e.g. P3 and P4) that are central to the core structure of group I introns. This implies a common function for the conserved T4 sequences. The only major structural differences among the T4 introns are the positions of the intron ORF and the lack of the P2 stem in the *nrdD* (sunY) intron. P2 is also an optional element in eukaryotic group I introns (Michel and Dujon, 1983; Waring and

Davies, 1984)

A significant degree of sequence conservation also exists beyond the intron boundaries extending 10-12 nucleotides to each side of the splice junction. This may reflect a target site for intron insertion and / or interactions with the core structure of the intron. Although many group I introns contain ORFs but they are looped out from the core structure. The position of the ORFs within the T4 introns (P6 and P9.1) is not typical. In addition, unlike many introns of this type , the T4 intron ORF overlaps structural elements and is thought to be involved in splicing (Hall et al, 1987). Both the *td* and *nrdB* intron ORFs extend past the P6 stem, ending within the highly conserved R sequence whereas the *sunY* intron has an ORF that overlaps P9.2, the last structural element of the intron. These arrangements are not without precedent. An intron in the large rRNA gene of *Chlamydmonas reinhardii* chloroplasts bears a remarkable structural resemblance to the *td* and *nrdB* introns sharing with them all of the unusual features (Rochaix et al, 1985). Similarly the placement of the ORF within the *nrdD* intron resembles that in the three fungal mitochondrial introns (Waring and Davies, 1984 ; Michel and Cumming, 1985). Although the sequences of the T4 introns (outside of the conserved group I elements) are not similar to their eukaryotic counterparts, the striking structural similarity between the prokaryotic and eukaryotic introns suggest a common evolutionary origin. These structural features are likely to reflect common functions carried out by RNA molecules from these widely divergent sources.

SELF SPLICING

In the genes of eukaryotes, coding sequences are often interrupted by stretches of noncoding DNA. The primary transcripts of such genes must undergo cleavage-ligation reactions. Two different mechanisms have been described by which catalytic RNAs can remove their own introns *in vitro*, and are classified as group I and group II self splicing introns. These introns are characterized by short conserved sequences as well as their potential to fold into complex secondary and tertiary structures that are conserved despite a high degree of primary sequence divergence (Jacquier, 1990). A third class of organelle introns the so called group III introns do not show self-splicing activity *in vitro* (Christopher and Hallick, 1989).

SELF SPLICING OF GROUP I INTRONS

Cech and colleagues (1981) found that a ribosomal precursor RNA of Tetrahymena could remove its own 413 nts long intron in the absence of proteins *in vitro*. This reaction proceeds by two consecutive trans-esterification reactions and requires a divalent cation (Mg^{+2} or Mn^{+2}) as well as guanosine (or a guanosine phosphate) which becomes covalently linked to the 5' end of the intron during the reaction. The 3'end of the 5' exon forms a hybrid with an internal guide sequence (IGS) in the 5' end of the intron. The trans-esterification reaction is initiated by the external guanosine, which is held in a guanosine binding site of the intron hydrogen bridges. The 3'-OH of this guanosine attacks the phosphorus atom at the 5' splice site and forms 3',5'-phosphodiester bond to the first nucleotide of the intron. Mg^{+2} is thought to stabilize the reaction intermediate by directly interacting with the 3' oxygen of the attacked phosphodiester bond (Piccirilli et al, 1993). Altered group I intron have been generated whose cleavage depends on the four fold larger Ca^{+2} ions (Lehman and Joyce, 1993). In the

second trans-esterification step, the free 3' hydroxyl group at the 3' splice site, brings about the ligation of the exons and removal of the linear intron (LIVS), the extra guanosine remaining at its 5' end. During this reaction, the conserved 3' guanosine from the intron remains bound at the G-binding site. Subsequently, the intron itself undergoes a self-catalyzed trans-esterification reaction. The conserved terminal 3' guanosine residue still bound in the G-binding site attacks a phosphorus atom near the end of the molecule resulting in intron cyclization (CIVS) and removal of short 5' oligonucleotide.

Group I self splicing is very sensitive to mutation in core of this structure but many of the peripheral stem-loop structures can be deleted without loss of splicing function *in vivo*. These structures may , however , play a role in stabilizing the intron or providing binding sites for proteins which facilitate or regulate self splicing *in vivo*. Introns themselves may infact encode a protein, involved in their own structural stabilization (Lambowitz, 1989).

SELF-SPLICING REACTION PATHWAY

Splicing: The self splicing of group I introns proceeds by two consecutive trans-esterification reactions. The first of these is initiated by guanosine or one of its 5'-phosphorylated forms (GMP, GDP, or GTP) which attacks the phosphorus atom at the 5' splice site and forms a 3',5'-phosphodiester bond to the first nucleotide of the intron. The 5' exon, now terminating in a free 3' hydroxyl group, then attacks the phosphorus atom at the 3' splice site (Fig. 3). This results in ligation of the exons and excision of the intron. This splicing pathway was first worked out for the *Tetrahymena thermophila* nuclear pre-rRNA (Zaug et al, 1983; Cech, 1987; Been et al, 1988). It was subse-

quently found applicable to self splicing of precursor RNA containing group I introns from diverse sources, including a *Neurospora crassa* mitochondrial pre-mRNA (Garriga and Lambowitz, 1984), several *Saccharomyces cerivisiae* mitochondrial pre-mRNA and pre-rRNA (Van der Horst and Tabak, 1985; Gampel and Tzagoloff, 1987; Partono and Lewin, 1988), several bacteriophage T4 and SPO1 pre-mRNA (Chu et al, 1986; Gott et al, 1986; Goodrich-Blair et al, 1989) and nuclear pre-mRNA from *Physarum polycephalum* and *Pneumocystis carinii* (Sogin and Edman, 1989).

Guanosine addition also occurs to the 5' end of the *N. crassa* mitochondrial rRNA group I intron during splicing *in vivo* (Garriga and Lambowitz, 1983). This RNA does not undergo detectible self splicing *in vitro*, but nevertheless shares this mechanistic hallmark of group I RNA splicing. The evidence for the reaction pathway shown in *N. crassa* includes characterization of the products and intermediates at the level of individual nucleotides and, in the case of Tetrahymena pre-rRNA, at individual phosphates. The phosphate to which G is added is derived from the UpA at the 5' splice site of the pre-rRNA (Zaug et al, 1983) whereas the phosphate at the ligation junction is derived from the GpU at the 3' splice site (Price et al, 1987a). The proposed intermediate has been characterized *in vitro* (Inoue et al, 1986) and observed in RNA produced *in vivo* (Kister and Eckert, 1987). Another prediction of the trans-esterification model that is the reversibility of the reaction in the absence of exogenous guanosine and the inversion of configuration around phosphorus have been verified. The nucleotidyl transfer reaction bears similarity to reactions catalyzed by DNA and RNA polymerases, with the difference that the nucleophile 3'-OH attacks a phosphorus atom involved in a phosphodiester linkage during splicing and one involved in a phosphoanhydride linkage during nucleic acid synthesis. These reactions are also similar to those catalyzed by DNA

topoisomerases and recombination enzymes. However, the nucleophile is a 3'-OH in splicing rather than an amino acid side chain (e.g. tyrosine hydroxyl group in the case of *E.coli* DNA gyrase; Tse et al, 1980). Two of the other three classes of RNA splicing also occur by trans-esterification with conservation of the number of O-P bonds throughout the reaction. These differ from group I RNA splicing in that the nucleophile is a 2'-OH instead of a 3'-OH.

Hydrolysis of Splice Site: Both the 5' and 3' splice site phosphodiester bonds of the Tetrahymena pre-rRNA are usually susceptible to hydrolysis (Zaug et al, 1984, 1985; Inoue et al, 1986). Thus in the absence of guanosine, the pre-rRNA is slowly cleaved at these sites. The cleavage products have 5'-phosphate and 3'-hydroxyl termini, consistent with the mechanism of splicing but opposite to the end-groups left by random alkaline hydrolysis of RNA. Site-specific hydrolysis is thought to reflect the ability of the folded RNA structure to activate the splice-site phosphates or the incoming nucleophile, using the same catalytic mechanism which is responsible for self splicing. The rate of hydrolysis of the phosphodiester bond following the 3'-terminal G of the intron is greatly accelarated; the second order rate constant is about 10 orders of magnitude higher than that calculated for hydrolysis of an average phosphodiester bond in RNA giving the same 5'-phosphate / 3'-hydroxyl end - groups (Zaug and Cech, 1986a, 1986b). Specific hydrolyses at the 5' splice site and at the cyclization site are also enhanced, although not as much as at the 3' splice site (Inoue et al, 1986; Tanner and Cech, 1987). Thus , while all phosphodiester bonds are energetically equal as far as their RNA synthesis is concerned, they can have vastly unequal stability in the context of a particular folded RNA structure (Cech, 1987).

Intron Cyclization and Trans-esterification Reaction: The trans-esterification mechanism was discovered by Blum et al (1991). The splicing of group I intron proceeds by two consecutive transesterification reactions . The first of these involves guanosine or one of its 5'-phosphorylated form (GMP, GDP or GTP) which attacks the phosphorus atom at the splice site and forms a 3',5'- phosphodiester bond to the first nucleotide of the intron (Fig. 3) . The 5'exon, now terminating in a free 3'-hydroxyl groups , then attacks the phosphorus atom at the 3'-splice site which results in ligation of the exon and removal of the intron (Herschlag and Cech, 1990a).

The evidence for this reaction pathway includes characterization of products and intermediates at the level of individual phosphate. The phosphate to which G is added is derived from the UpA at the 5' splice site of the pre-rRNA (Cech, 1987). The phosphate at the ligation junction is derived from the GpU at the 3' splice site (Price et al, 1987b). The divalent cation (Mg^{+2} or Mn^{+2}) are also required for efficient splicing (Menger et al, 1996). The Tetrahymena intron undergoes an intra molecular cyclization reaction after the removal from pre-rRNA (Garriga and Lambowitz, 1983). This reaction is also self catalyzed and again occurs by transesterification, with the 3'-terminal G-OH of the RNA attacking of phosphorus atom located near the 5' end of molecule (Zaug et al, 1983). A 3' terminal guanosine is conserved among group I introns and cyclization is common but not universal among this group.

A variety of self catalyzed, splicing related reactions including cyclization of excised IVS RNA, oligomerization of IVS RNA and intramolecular exon ligation were found to take place by concomitant cleavage and rejoining of RNA (Inoue et al, 1986; Tabak et al, 1987). These were readily explained by trans-esterification model. The postulated *splicing in-*

termediate was subsequently detected, purified and found to undergo second step of splicing on incubation with an oligonucleotides analog of 5' exon (Zaug and Cech, 1985). The splice junction phosphate was derived from the 3' splice site as predicted. Finally, a transesterification reaction and the cyclization of excised IVS RNA have been shown to be reversible (Sullivan and Cech, 1985).

SPLICE SITE

There is a simple rule that determines the 5' splice site of all group I introns that is the last nucleotides of the 5' exon are involved in base pairing (called P1) , with a complementary sequence located within the intron, 5' to its active core. In almost all cases (including *nrdD*) the exon ends with a U, which was paired to a G in the intron. This rule which was originally proposed as the basis of sequence comparisons (Davies et al, 1982; Michel et al, 1982) has subsequently been experimentally confirmed (Been and Cech, 1985, 1986; Waring et al, 1986). The P1 pairing may be sufficient to specify the 5' cleavage site because point mutations in the exon destabilizing P1 can activate cryptic 5' splice sites that mimic the original pairing (Chandry and Belfort, 1987; Price et al, 1987b). The specificity at the 3' splice site is not well understood. All group I introns end with a G and changing that nucleotide to U or C in the intron of the large rRNA of *Tetrahymena thermophila* allows only residual splicing (Price and Cech, 1988). Also the two nucleotides immediately preceding the 3'- terminal G of the Tetrahymena intron have been shown to contribute significantly to binding of the 3' splice junction by the catalytic core (Tanner and Cech, 1987). Indeed mutation in the penultimate residues of the intron results in slowing the rate of splicing, without changing its specificity (Price and Cech, 1988) . Moreover, Davies et al (1982) noted that 5' to the intron residues that pair with the 5' exon to

form helix P1, there are usually several additional residues that can pair (P10) with the first several residues of the 3' exon. Most group I introns would thus include an internal guide sequence which, by pairing with both 5' and 3'exon residues (P1 and P10) would align the two exons for the ligation step (Shub et al, 1988; Doudona and Szostak, 1989a, 1989b; Xu and Shub, 1989).

T4 INTRONS

The bacteriophage T4 genome contains three introns (Fig. 4), at least two of which are in genes coding for enzymes involved in nucleotide biosynthesis. The first T4 intron was discovered in 1984 in the thymidylate synthase (*td*) gene (Chu et al, 1984). It raised the question of generality of introns and RNA splicing in prokaryotes. While introns have been reported to exist in archae-bacterial tRNA (Kaine et al, 1983; Daniel et al, 1985) and rRNA (Kjems and Garrett, 1985), the phylogenetic status of these prokaryotes is unusual. Archaebacteria have more characteristics in common with eukaryotes than do eubacteria and may form a third primary kingdom (Fox et al, 1980). This 1016 bp intervening sequence in *td* gene of T4 splits into two exons of 183 amino acids (exon I) and 103 amino acids (exon II). Subsequently, an intron of 598 bp was found on the *nrdB* gene encoding the small subunit of ribonucleotide reductase (Sjoberg et al, 1986) and an intron of 1030 was found in *nrdD* (previously abbreviated as sunY) gene encoding an anaerobic ribonucleotide reductase (Gott et al, 1986).

SPLICING OF T4 INTRONS

The T4 introns have the ability to self-splice *in vitro* (Chu et al, 1986; Gott et al, 1986). Autocatalysis is enabled by folding of the intron into characteristic secondary structure that facilitates a series of trans-esterification reactions necessary for intron excision and exon ligation (Belfort, 1990). This critical ribozyme secondary structure comprises of ten RNA pairing regions (P1- P10) which along with conserved primary sequence elements P,Q,R,S identify the T4 intervening sequences as group I intron (Cech, 1988; Shub et al, 1988; Saldanha et al, 1993). The T4 *td* intron contains four long range base paired regions in P3, P6, P7 and P10 . The remaining base paired regions are stem-loop structure including P7.1 and P7.2 which coupled with specific variants in the P,Q,R,S sequence elements further subclassify the T4 introns as group IA (Cech, 1988; Belfort, 1990).

The secondary structure model for group I intron was initially postulated on the basis of phylogenetic comparisons of intron sequences from various organisms which predicted a common secondary structure (Davies et al, 1982; Michel et al, 1982; Waring and Davies, 1984). Since then, mutational analysis of autocatalytic group I intron in Tetrahymena, T4 and yeast mitochondria has identified many splicing defective cis-acting intron mutations that would destabilize predicted RNA pairing regions , thus generally supporting the model. In T4, for example, the sequence changes for 31 splicing defective *td* intron mutants (19 phage mutants, 12 mutations from a cloned *td* gene) have been determined and the vast majority disrupt proposed intron RNA helices (Chandry and Belfort, 1987; Hall et al, 1987; Ehrenman et al, 1989; Brown et al, 1993).

Single base changes in predicted RNA pairing regions which disrupt

splicing, provides proof that a nucleotide is essential for the splicing reaction. Genetic verification that a base pairs (and thus a pairing region) exists, can be demonstrated by the isolation of a pseudorevertant that harbours a compensatory mutation at the nucleotide predicted to base pair with the original mutant nucleotide. The double mutant would be predicted to regenerate a Watson-Crick base pair in an essential RNA helix, thus permitting intron RNA to fold, resulting in the restoration of some degree of autocatalytic splicing. Such experiments have been carried out using site-specific mutagenesis to verify the existence of group I intron RNA pairing region in Tetrahymena and yeast mitochondria (Weiss-Brummer et al, 1983; Holl et al, 1985; Burke, 1988; Flor et al, 1989; Williamson et al, 1989; Couture et al, 1990). For the T4 introns, however, only the P6 pairing region of the *td* intron has been proven by isolation of such second site suppressor mutations (Ehrenman et al, 1989).

PROBABLE ROLE OF INTRONS IN T4 PHAGE

It is generally true that prokaryotic genes do not have introns. Because prokaryotic genomes tend to be smaller and replicate faster than eukaryotic genomes, it is thought that introns and other non-essential DNA sequences have been kept out of prokaryotes by selective pressure to streamline the genome (Hickey et al, 1989). An extreme exception is the occurrence of three group I introns in the coliphage T4 (Belfort, 1990). Phage T4 has a 166 Kb DNA genome, which otherwise appears to be highly streamlined, for example, promoters and even translational initiation regions are often found to lie nested within the coding sequence of upstream genes. The presence of these introns in the space-efficient T4 genome suggest that there is some selective pressure keeping them there.

The introns occur in the genes encoding the phage dTMP synthase (*td*), the small subunit of the phage nucleotide reductase (*nrdB*), and an anaerobic ribonucleotide reductase (*nrdD*). It is curious that the two proteins of a known functions catalyse the steps in the same pathway i.e. *de novo* dNTP synthesis from rNTP pool. Moreover, they lie at the only two steps that consume reducing equivalents (ultimately NADPH) from the reserves of the host cell. Group I splicing itself relies on an exogenous ribonucleotide and may be inhibited by guanosine analog. These introns are involved in a feedback regulatory loop that controls the level of dNTP synthesis in the infected cell and, therefore, confer a selective advantage upon intron- containing phage (Gott et al, 1986; Goodrich-Blair et al, 1990).

A different hypothesis proposes that the introns are mobile genetic elements. This was suggested by the inconsistent occurrence of the introns in the closely related phages T2 and T6 (T2 has no intron, and T6 has only *td* intron) (Pedersen-Lane & Belfort, 1987; Quirk et al, 1989a). This hypothesis has been substantiated by recent direct experimental evidence demonstrating that the *td* and *nrdD* introns are indeed actively mobile (Quirk et al, 1989b) and both introns encode DNA endonucleases highly specific for intron-less *td* (Bell-Pedersen et al, 1989, 1990; Chu et al, 1990) or *nrdD* genes (Bell-Pedersen, 1990). The resulting double-stranded DNA breaks appear to initiate an efficient unidirectional gene conversion event, resulting in the conversion of intron-less to intron-plus gene in a mixed infection. This high-frequency conversion to intron plus genes might be expected to easily maintain the introns in the phage population, even in the face of other selective pressure to remove them. Group I introns mobility appears to be general phenomenon (Dujon, 1989; Perlman and Butow, 1989; Scazzacchio, 1989). Although high frequency intron mobility may account for the presence of the *td* and *nrdD* introns in T4, it fails to explain the

presence of the immobile *nrdB* intron.

SPLICING OF INTRONS IN PROKARYOTES

Splicing of introns in prokaryotes are carried out by two trans-esterification reactions initiated by nucleophilic attack of guanosine at 5' splice site. This process depends on conserved secondary and tertiary structures that direct folding of the intron such that the 5' and 3' splice sites are juxtaposed to the guanosine-binding site within the introns' catalytic core (Cech, 1990; Michel and Westhof, 1990; Cech et al, 1992; Saldanha et al, 1993). The Tetrahymena intron splices 50-fold more efficiently *in vivo* than *in vitro* (Brehm and Cech, 1983) suggesting for the involvement of accessory factors during *in vivo* splicing. In addition, two groups of accessory proteins that regulate splicing of fungal mitochondrial introns have been defined by genetic criteria (Burke, 1988; Lambowitz and Perlman, 1990). First type is the introns encoded proteins called maturase that facilitate splicing of their cognate introns (Anziano et al, 1982; De La Salle et al, 1982; Lazowska et al, 1989). Second type belongs to the nuclear-encoded splicing proteins (Burke, 1988; Lambowitz and Perlman, 1990). The best characterized of these proteins are found in Neurospora where a mutation in the *cyt*18 gene blocks splicing of a number of group I introns *in vivo*, including the COB1 intron, which can self splice *in vitro* (Garriga and Lambowitz, 1984, 1986). CYT-18 protein which corresponds to mitochondrial tyrosyl tRNA synthetase (Akins and Lambowitz, 1987) binds specifically to the mitochondrial large subunit rRNA (LSU) intron with high affinity (Guo and Lambowitz, 1992). CYT-18 can also promote splicing of mutant T4 *td* introns both *in vivo* and *in vitro* (Mohr et al, 1992) suggesting that the protein acts by stabilizing the correct tertiary structure of the introns (Guo and Lambowitz, 1992).

A role for protein factors in modulating splicing of the T4 introns *in vivo* is suggested by several observations. First, both the *nrdB* and *nrdD* introns require elevated monovalent ion concentrations to splice efficiently *in vitro* (Hicke et al, 1989). Second, utilization of a cryptic 5' splice site by the *td* intron is reduced *in vivo* compared with *in vitro* (Chandry, 1991). Third, the *td* intron splices more efficiently *in vitro* at elevated temperature (Chu et al, 1987). Fourth, analysis of a *td* trans-splicing reaction revealed that the *in vitro* reaction occurs efficiently only at high temperature (55° C), whereas the *td* intron is capable of trans-splicing *in vivo* at physiological temperature (Galloway-Salvo et al, 1990). Several findings suggested that facilitatory factors may be present within the cell, prompting attempts to purify these putative factors. Using stimulation of trans-splicing as an initial *in vitro* assay, Coetzee et al (1994) identified an activity that facilitated both trans- and cis-splicing of the *td* intron. This activity, which also promoted splicing of the *nrdB* and *sunY* (*nrdD*) introns, comprises multiple proteins that have different activities for splicing enhancement. Several of these are ribosomal proteins, with ribosomal protein S12 being the most active (Coetzee et al, 1994). Functional analysis suggests that these proteins are mechanistically distinct from those that specifically enhance group I splicing and, rather, serve to facilitate formation of the active conformation of the intron by acting as RNA chaperones (Coetzee et al, 1994).

CHARACTERISTICS OF T4 *td* INTRON

It has been reported that the *td* gene encoding T4 phage thymidylate synthase is interrupted by a 1017 base pair intron between the triplets coding for glycine 183 (GGT)and leucine 184 (CTA) (Chu et al, 1984). Expression of *td* gene was shown subsequently to involve post transcriptional

processing of the primary transcript both *in vivo* (Belfort et al, 1985) and *in vitro* systems (Chu et al, 1985). The nucleotide sequence of the entire *td* intron has now been determined using overlapping deletions of the *td* gene in M13 recombinant clones. As in the case of most class I introns, which include the self splicing intron of the nuclear large rRNA in *Tetrahymena thermophila* and the introns in yeast and fungal mitochondrial genes (Cech, 1983) the *td* intron is preceded by a U and ends in a G. In addition, the nucleotide sequences at the *td* splice sites of exonI-intron (GGUUAA) and intron-exonII (AUGCUG) show a fair degree of homology to their eukaryotic counterparts. Furthermore, at least four conserved nucleotide sequences (boxes P,Q,9L,and 2) found in class I introns (Burke and Rajbhandary, 1982; Michel et al, 1982; Waring and Davies, 1984) were discernable in the *td* intron and exhibited some degree of homology. Moreover, the box 9L and 2 sequences are highly conserved in the *td* intron especially the internal 5 nucleotide stretch in each box. These pentanucleotide sequences have been shown to form a stable base-paired region in the Tetrahymena rRNA that is believed to facilitate the proper intron excision and exon splicing from precursor RNA (Burke et al, 1986, 1987).

SPLICING OF *td* INTRON

The splicing in *td* intron involves covalent addition of a guanosine residue to its 5' end by an apparently energy-independent trans-esterification reaction (Hall et al, 1987). One end product of the reaction is a circular intron RNA that is likely to be covalently closed by a 3'-5' linkage, with concomitant loss of the two 5' residues of the intervening sequence (Hall et al, 1987). Loss of nucleotides from the 5' intron terminus suggests that this cyclization event is mechanistically similar to that of the Tetrahymena rRNA intron (Zaug et al, 1983) and the yeast mitochondrial rRNA intervening se-

quence (Tabak et al, 1984). In these cases, 15 and 4 nucleotides, respectively are removed upon formation of the 3'-5' linkage between the 3'terminal guanosine and the selected internal intron residue. In addition to cyclization site between the 15 and 16 nucleotide of Tetrahymena intron, there is minor site between the residues 19 and 20 (Zaug et al, 1984). Infact the 3' guanosine adds to an internal bond, 3' to a uridine residue. This observation suggests that the conserved uridine may be a major determinant in cyclization site recognition, much as the conserved uridine at the 3' end of the upstream exon is important in targeting the 5'splice site. Davies and co-workers (1982) proposed pairings between internal guide sequences and exon sequences around the splice sites as major determinants in targeting splice sites. Furthermore, since di- and tri-nucleotides similar to the sequences preceding the Tetrahymena rRNA 5' splice site (UCU) and the major and minor cyclization sites (UUU and CCU) are capable of reopening the circle, Sullivan and Cech (1985) proposed a trinucleotide binding site within the intron that guides the choice of the appropriate cleavage-ligation site. Consistent with such a model is the similarity of the sequences upstream of the yeast mitochondrial rRNA 5' splice site (GAU) and cyclization site (AAU). In case of the *td* gene there are only 2 nucleotides, (i.e. GU) preceding the cyclization site, the guanosine being the non-coded adduct. These residues are, however, identical to the GU upstream of 5'splice site. It would appear, therefore, that in *td* RNA splicing a dinucleotide may be sufficient to impart specificity to splice site selection. Regardless of evolutionary relationship, the parallels between *td* intron and other group I intervening sequences make the *td* gene an attractive model system for studying this type of reaction, using facile prokaryotic genetic techniques (Ehrenman et al, 1986).

THE *nrdB* GENE OF T4 PHAGE

Bacteriophage T4-encoded ribonucleotide reductase is required in the assembly of the dNTP synthetase complex (Chiu et al, 1982; Moen et al, 1988) and appears to be the last factor and the keystone in the assembly of the complex (Hilfinger and Ping, 1994). The alpha and beta chains of this alpha 2 beta 2 enzymes are encoded by the T4 *nrd*A gene and intron containing *nrdB* gene, respectively (Tseng et al, 1988, 1990). The synthesis of the beta 2 subunit is initiated after the synthesis of alpha 2 protein begins (Tomich et al, 1974; Chiu et al, 1976, 1982).

The *nrdB* intron of bacteriophage T4 contains 598 base pairs. This intron is closely related to other group I introns of T4 and eukaryotes (Sjoberg et al, 1986). The secondary structure of *nrdB* intron like *td* and other group I introns is comprised of various stem-loop structures. Besides these stem-loop structures, other pairing regions are also found to facilitate the secondary structure formation of intron (Davies et al, 1982; Cech et al, 1983; Michel and Dujon, 1983). The *nrdB* intron belongs to a sub-group, IA, that is characterized by systematic variations in the highly conserved sequences P, Q, R, and S (Michel and Dujon, 1983; Waring and Davies, 1984). It contains an extra domain (stem-loop P7.1 and P7.2) between the 3' end of P7 [5'] and 5' end P3 [3'] (Waring and Davies, 1984). The P6 pairing of the *nrdB* intron consists of 3 base pairs, two of which are G-C pairs and one is a G-U pair. Sequences peripheral to P6 are not required for *td* intron splicing. Infact all the splicing-defective mutants in the *td* intron clustered outside the ORF (Belfort et al, 1987; Ehrenman et al, 1989). Compensatory mutants in *td* exhibit only partial function *in vivo* and are inactive under stringent *in vitro* conditions, demonstrating that although the P6 pairing is necessary for splicing activity, there is a sequence specificity for P6 [5'] and / or P6 [3']. These functional constraints appear to be reflected in the

high degree of conservation of P6 among all two 3' residues of element Q and P6 [3'] forming the two 5' residues of R (Q and R being two of the four conserved group I sequence elements (Lal and Hall, 1997).

ETHYLMETHANE SULPHONATE (EMS) AS A MUTAGEN

EMS is an alkylating agent which was discovered by Loveless to induce mutations in Bacteriophage T2 (Loveless, 1958). The induced mutation may be delayed by several genarations (Green and Krieg, 1961) . This mutagen generally reacts with three (i.e. A, G, and C) of the four bases of DNA (Brookes and Lawley, 1960; Pal, 1962). EMS induces the GC to AT at a much higher rate than the AT to GC transition. The replacement of purine-pyrimidine pair by a pyrimidine-purine pair may also be induced to some extent by EMS (Freese, 1961; Freese and Freese, 1966)).

MOLECULAR MECHANISM OF EMS-INDUCED MUTATION

EMS reacts with guanine, adenine, cytosine and 5-hydroxymethyl-cytosine when the free bases or nucleotides are exposed in solution (Brookes and Lawley, 1960, 1962; Pal, 1962). The reaction of this agent with DNA produces 7-ethylguanine and 3-ethyladenine and other products may be formed to a lesser extent. Two distinct mutagenic mechanisms have been postulated for 7-ethylguanine formation by EMS and one mechanism for 3-ethyladenine (Pal, 1962).

1. *GC To AT transition mediated by ethylguanine pairing errors:* The guanine has a pk of 9.2 for the dissociation of the hydrogen at the N_1 position, whereas the corresponding pk for a 7,9-substituted guanine is shifted closer to 7 (Lawley and Brookes 1961). Thus, ionization of the N_1 may occur while

the DNA strand is serving as a template for DNA synthesis which would permit pairing of 7-ethylguanine with thymine (Fig. 5A) as well as with cytosine. This would generate a GC to AT transition. If a template molecule of 7-ethylguanine was not ionized at the moment of DNA replication it would pair like guanine with cytosine and hence make a normal, non-mutated copy.

2. *AT TO GC transition mediated by 3-ethyladenine:* 3-Methyladinosine has been shown to have an imino group in place of the amino group of adenine and to be positively charged at pH 8 and below (Pal, 1962). The presumed structure indicates that during DNA replication, pairing would be possible between 3-ethyladenine and cytosine by two hydrogen bonds (Fig. 5B) and that pairing with thymine would not occur.

SCOPE OF THE WORK

The group I introns are subjected to undergo self-splicing reaction which is close to the action of ribozyme (Cech, 1987; Michel and Westhof, 1990). The studies on the T4 introns have also revealed, that the mode of catalytic action of these introns are similar to that of other group I introns (Hall et al, 1987; Lal and Hall, 1993, 1997). Ribozymes are small RNA molecules endowed with endonuclease activity at enzymatic rate (Symons, 1992, 1994; Tusch et al, 1995). Upon formation of a complex with their complementary target RNA , cleavage takes place leading to degradation of the target , on the other hand, antisense oligodeoxyribonucleotides similarly complex with RNA but, instead, inhibit translation (Fig. 6). Tetrahymena ribozyme is well characterized by P4-P6 structural domain which is an independently folding domain of intron tertiary structure (Wang et al, 1993). This P4-P6 domain was measured as 2.8 A° crystal structure (Cate et al, 1996).

Six naturally occurring ribozyme structures have been identified. Cech (1987) described the Tetrahymena ribozyme which belongs to the Tetrahymena group I intron. This molecule is capable of autocatalytic cleavage during the RNA splicing process in Tetrahymena. A second ribozyme structure was elaborated by Guerrier-Takada et al (1983) during experiments which characterized the structure of ribonuclease P. RNAse P is an endonuclease capable of cleaving the 5' end of pre-transfer RNA molecules during the maturation process leading to the formation of functional tRNA. Of the two subunits which comprise this molecule, M1-RNA bears the catalytic properties. Augmentation of the endonuclease activity appears to be the function of the 14 KD second subunit. A third type of ribozyme was identified in the minus-strand of satellite RNA of the tobacco ringspot virus (Hampel and Tritz, 1989; Hampel et al, 1990; Mitchell et al, 1992). Owing to the characteristic hairpin like structure formed upon hybridization of the cataytic domain, consisting of 50 nucleotides with the substrate strand, it is referred to as hairpin ribozyme. Hammerhead ribozymes (Uhlenbeck, 1987), a fourth class of ribozymes , have a common secondary structure consisting of 3 stem structures flanking a conserved single-stranded region which possesses catalytic activity. It has been identified in plant viroids, virusoids and satellite viruses. Two additional catalytically active ribozyme motifs have been described in hepatitis delta virus and in Neurospora (Xu and Shub, 1989).

This enzymatic activity of group I introns and ribozyme heralded a new era in the field of modern medicine. The natural activity of ribozymes involved a single intramolecular cleavage reaction designated as cis-reaction because of the presence of the both enzyme and substrate within the same molecule. By separating the hammerhead ribozymes through the re-

moval of naturally occurring loop in the stem structures, Forster and his associates as well as Uhlenbeck in 1987 have demonstrated trans reaction which allows for the turn over of the several substrate molecules. The control of expression of specific gene is an application of ribozyme technology which is of particular interest. Because of its small size, the catalytic core of hammerhead ribozyme can be integrated into RNA or DNA structures and subsequently cleave their target.

This catalytic activity of ribozyme is being exploited for the inhibition of the replication of human immunodeficiency virus (HIV) through the inactivation of viral RNA with ribozymes (Sarvar et al, 1990; Dropulic et al, 1992). Besides that, the hairpin and hammerhead ribozymes containing a 5'C(UUCG)G3' loop were designed to cleave the long terminal repeat of HIV-I (Koizumi et al, 1995). As it is an established fact that hepatitis C virus (HCV), a positive strand RNA virus is the major infectious agent responsible for causing chronic hepatitis. Currently, there is no vaccine for HCV infection. To investigate the utility of the ribozyme technology for the management of HCV infection, six hammerhead ribozymes directed against a conserved region of the plus strand and minus strand of HCV genome were isolated. The expressed ribozyme individually or in combination were efficient at reducing or eliminating the respective plus- or minus- strand of HCV RNA expressed in culture cells and from chronic HCV infected patients (Lieber et al, 1996).

Ribozyme is thought to be a powerful tool for the suppression of specific gene expression, but there are many failures under *in vivo* conditions. One reason may be that there are many cellular proteins *in vivo* which may inhibit the catalytic activity of ribozymes (Koguma et al, 1995; Junn and Kang, 1996). Therefore, to achieve a successful trial *in vivo* a chemical

synthesis of RNA has been designed which has several advantages over synthesis by "atomic mutagenesis" in which individual atoms or chemical groups are altered or removed by the site-specific incorporation of nucleotide analogs. Advances in the solid-phase chemical synthesis of RNA have allowed the complete chemical synthesis of small ribozymes, including 19- and 35- nts hammerhead ribozyme (Perrault et al, 1990; Scaringe et al, 1990) and 59-nts hairpin ribozyme (Vinayak et al, 1992). The small sized synthetic ribozymes are easy to deliver through liposome technology.

These developments in the ribozyme technology may be challenging to the scientists working on the synthesis of vaccines against HIV and other RNA viruses.

DETAILED OBJECTIVES OF THE WORK

In view of the available literature, we initiated the work on the screening of T4 *nrdB* mutants to study the structural functional relationship in the RNA splicing of group I introns as a model to understand the requirements of secondary and tertiary structure of RNA in the autocatalytic cleavage of introns. The focus of the research was, therefore, to address the problem on the following lines:

(1) Isolation and screening of the EMS-induced *frd*1 double mutants by using a novel technique of halo phenotype.

(2) Mapping of *frd*1*nrdB* mutants with the help of recombination based marker rescue technique with various clones and sub-clones carrying the whole *nrdB* gene, *nrdB* intron, as well as carrying various subregions of intron.

(3) Dot blot hybridization splicing assay of the RNA preparations of the mutants by using various probes to analyse the splicing defective point

mutations.

(4) EMS-mutagenesis of *frd*1*nrdB* double mutants.

(5) Isolation and mapping of intragenic and extragenic revertants of *frd*1*nrdB* mutants.

(6) Dot blot analysis of those revertants.

(7) Elucidation of secondary and tertiary structure responsible for RNA splicing taking advantage of the characteristics of the splicing defective mutants and their revertants.

Figure 1: Splicing mechanism of the four major groups of precursor RNAs. Wavy lines indicate intron, smooth lines indicate flanking exons. For nuclear mRNA splicing, many components assemble with the pre-mRNA to form the spliceosome; only two, the U1 and U2 small nuclear ribonucleoproteins, are shown [Cech, 1990].

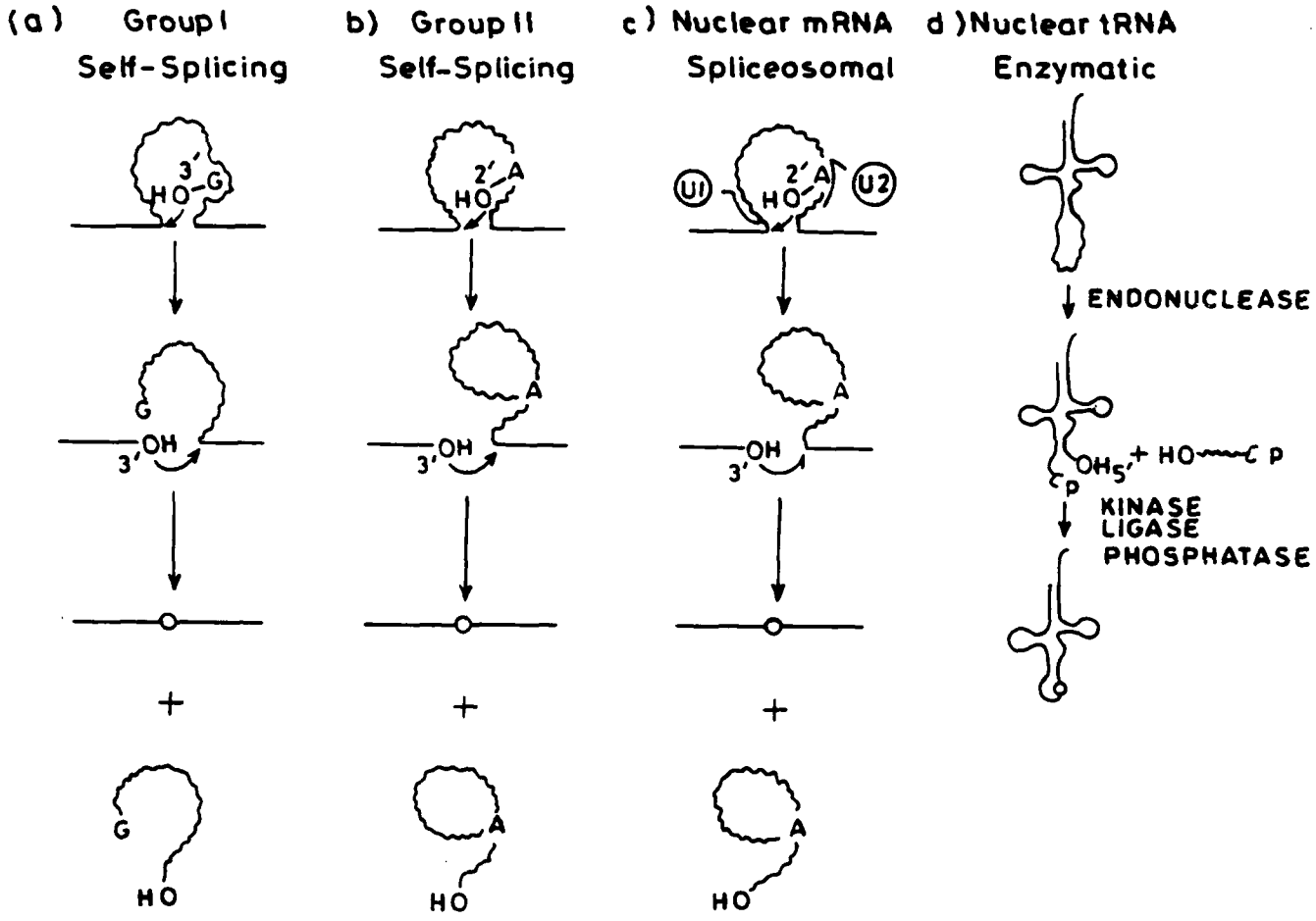

(a) Group I
Self-Splicing
b) Group II
Self-Splicing
c) Nuclear mRNA
Spliceosomal
d) Nuclear tRNA
Enzymatic
3'
HO
G
2'
A
U1
U2
OH
ENDONUCLEASE
OH 5'
+ HO
p
KINASE
LIGASE
PHOSPHATASE

Figure 2: A Typical structure of group I intron, according to the convention of Burke et al (1987). (a) Conserved sequence elements P, Q, R, and S are represented by their most common nucleotide sequences. Dashed line between P4 and P6 was added to make the diagram less crowded; no nucleotides have been omitted here. Filled arrows, 5' and 3' splice sites; Open arrow, site of insertion of extra stem-loop(s) P7.1 and P7.2 in group IA intron. (b) Secondary structure of *T. thermophila* intron with base-paired regions.

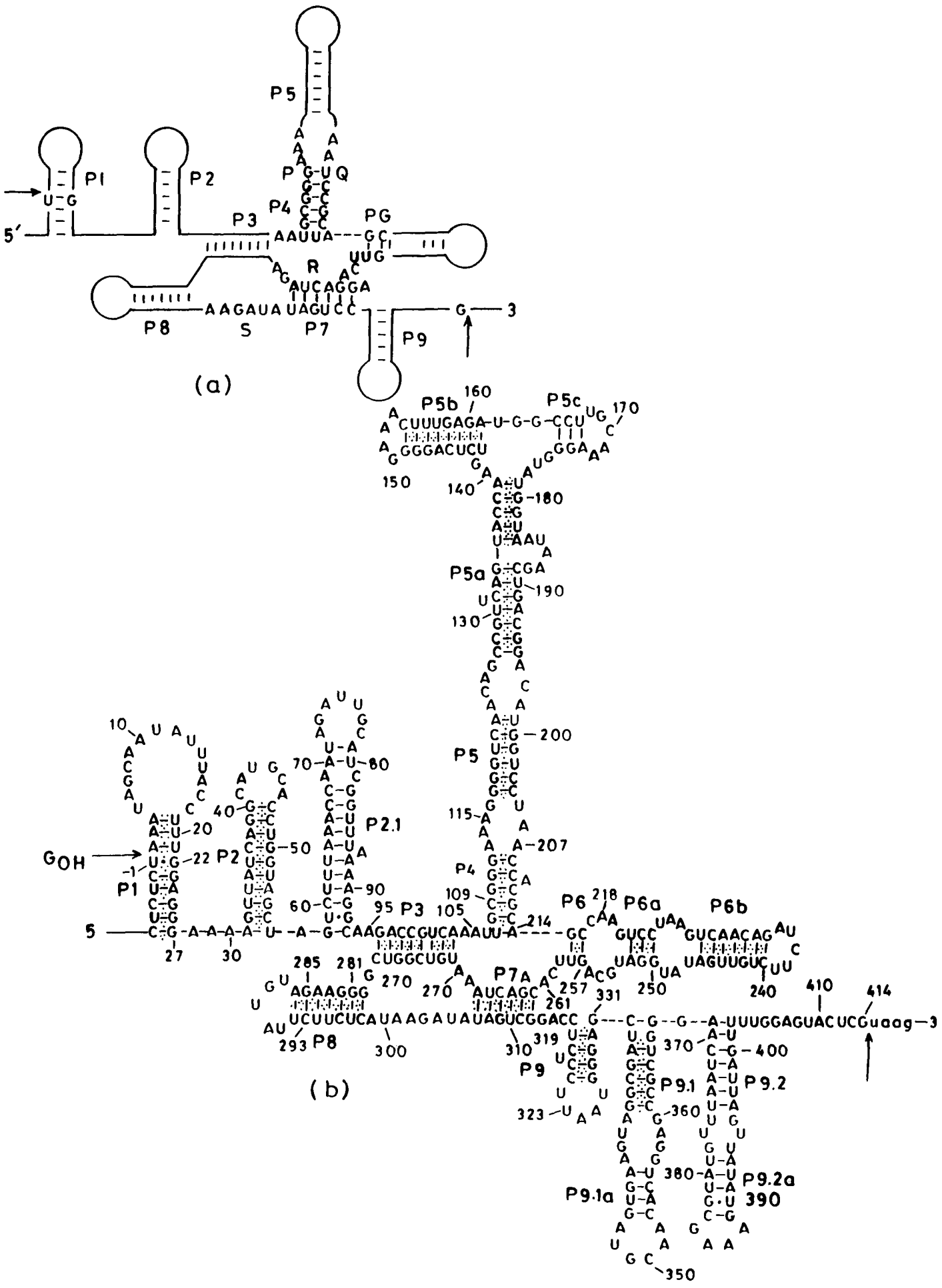
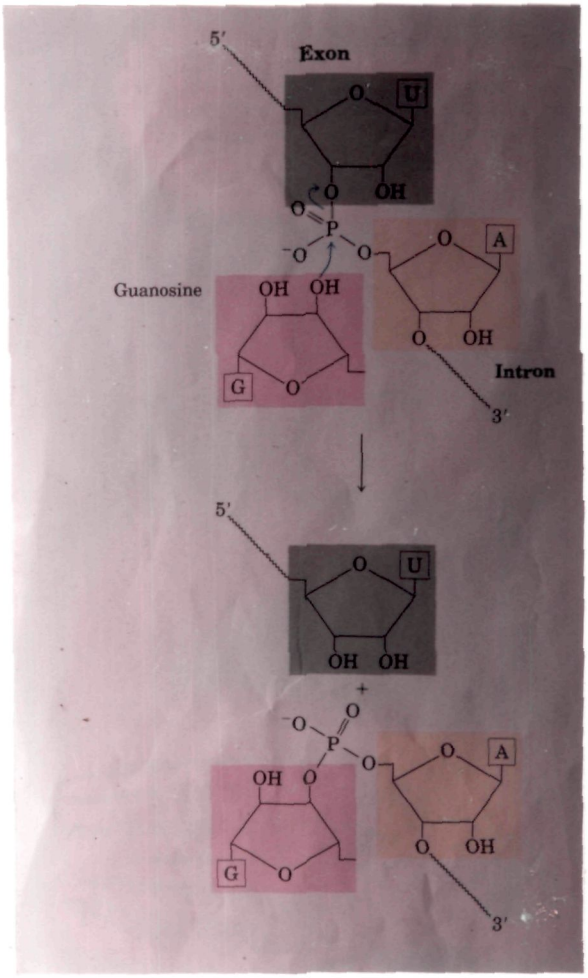

P5
P1
P2
P3
P4
P8
S
P7
P9
5′
3
(a)
P5b
P5c
P5a
P5
P2.1
P6
P6a
P6b
P7
P9
P9.1
P9.2
P9.1a
P9.2a
GOH
(b)

Figure 3: Splicing mechanism of group I introns, and trans-esterification reaction, the first step in the splicing of group I introns in which the 3' OH of guanosine molecule acts as nucleophile.


5′
Exon
U
OH
Guanosine
OH OH
G
A
OH
Intron
3′
5′
U
OH OH
A
OH
3′

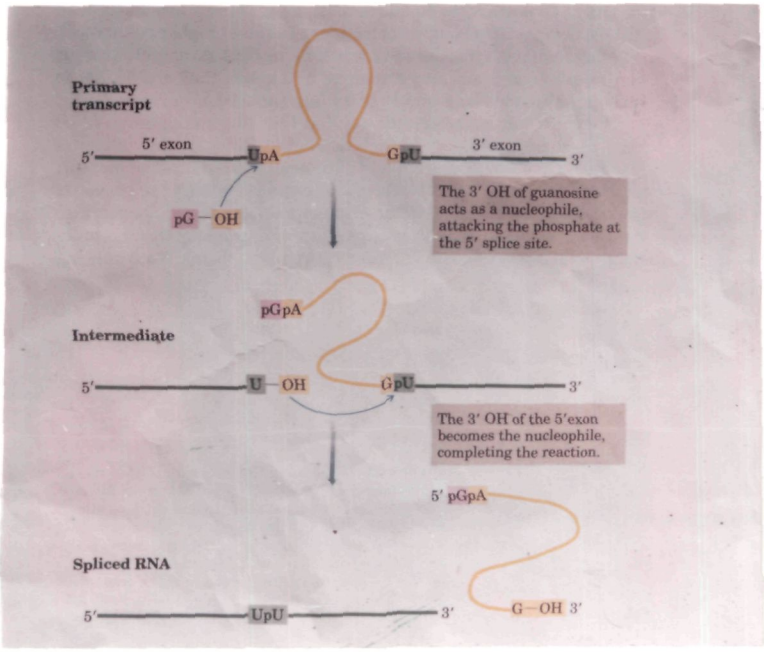
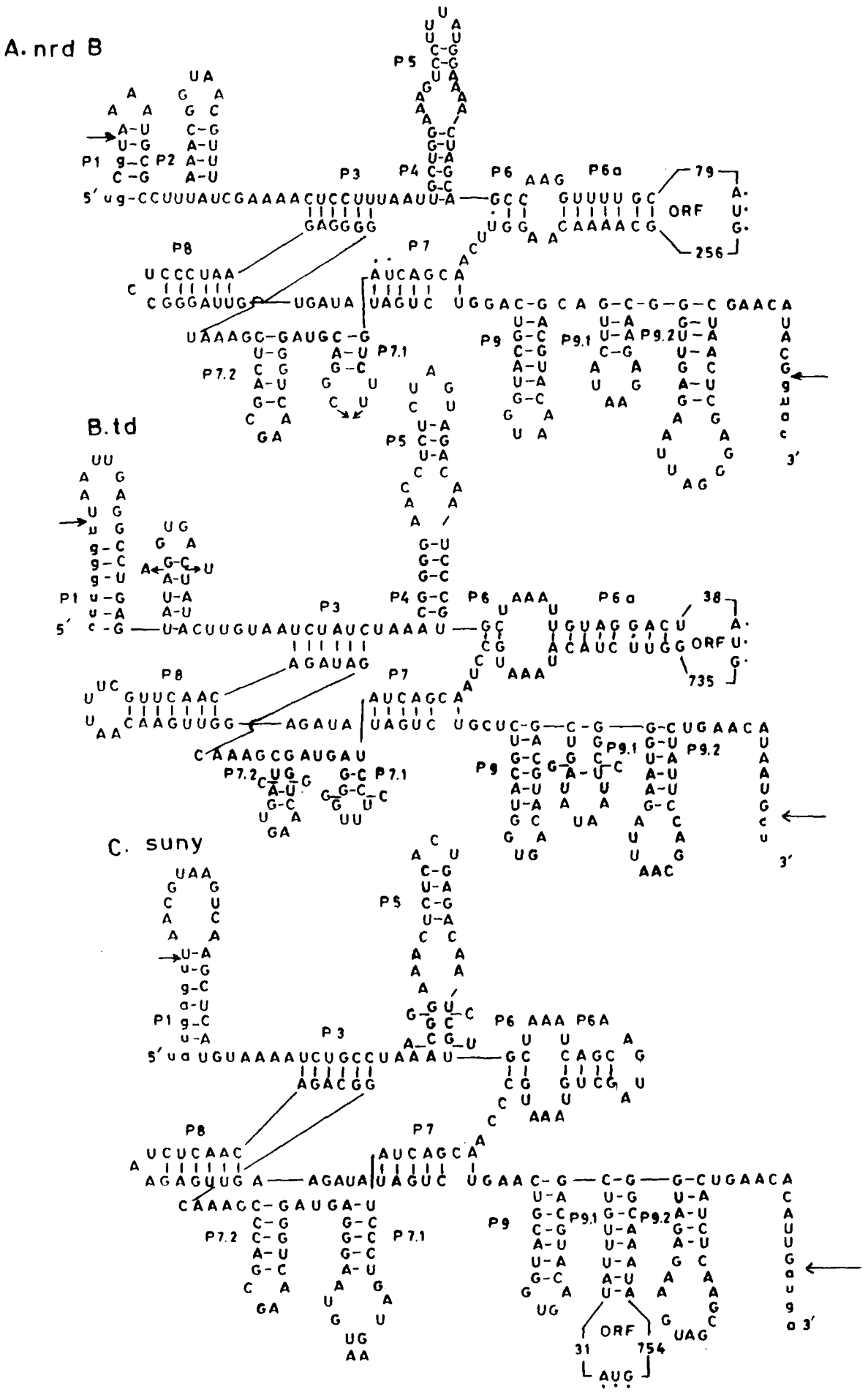

Primary transcript
5′ exon
5′
UpA
GpU
3′ exon
3′
pG—OH
The 3′ OH of guanosine acts as a nucleophile, attacking the phosphate at the 5′ splice site.
pGpA
Intermediate
5′
U—OH
GpU
3′
The 3′ OH of the 5′exon becomes the nucleophile, completing the reaction.
5′ pGpA
Spliced RNA
5′
UpU
3′
G—OH 3′

Figure 4: Secondary structures for intron sequences from the *nrdB*, *td* and *SunY* genes. (waring and Davies, 1984).


A. nrd B
B. td
C. suny
ORF
P1
P2
P3
P4
P5
P6
P6a
P7
P7.1
P7.2
P8
P9
P9.1
P9.2
AUG
UAG

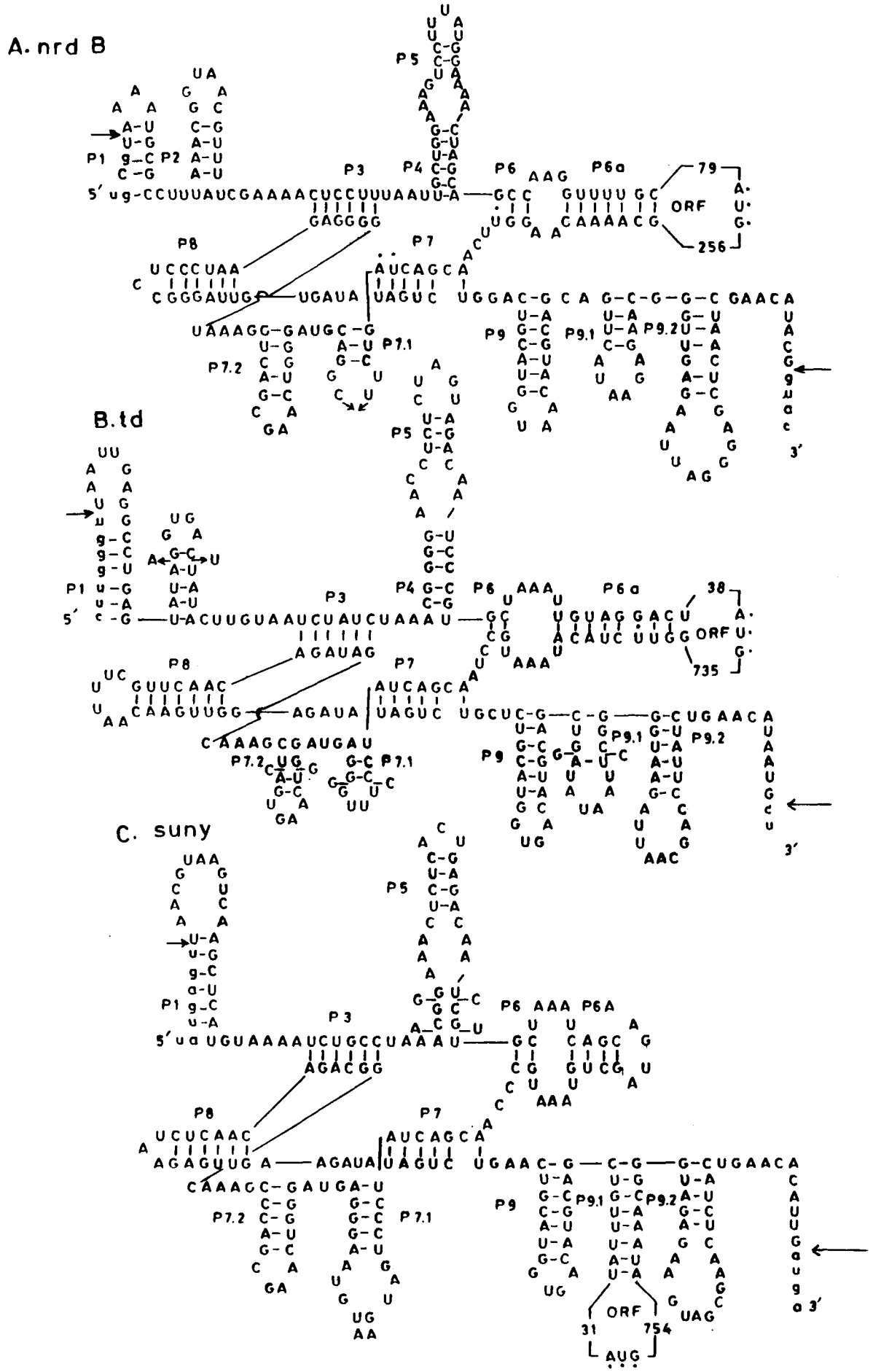
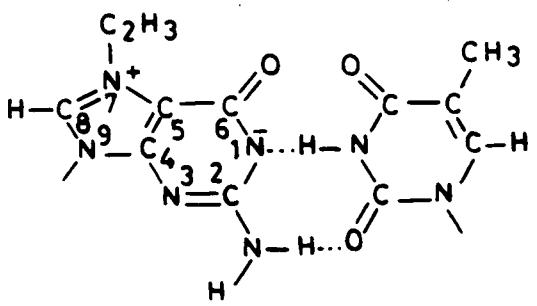

A. nrd B
B. td
C. suny
P1
P2
P3
P4
P5
P6
P6a
P7
P7.1
P7.2
P8
P9
P9.1
P9.2
ORF
5'
3'
79
256
38
735
31
754
AUG
UAG

Figure 5: Pairings postulated to occur during replication of DNA containing ethylated purines. Hydrogen bonds are shown by dotted lines. A, 7-ethylguanine and thymine; B, 3-ethyladenine and cytosine.

(A)



(B)

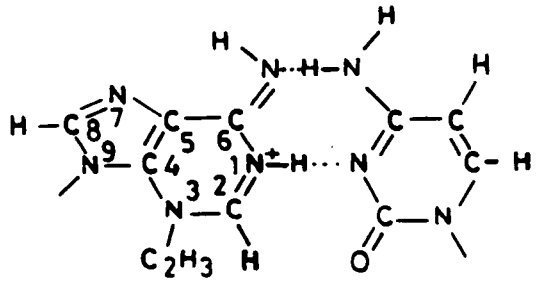
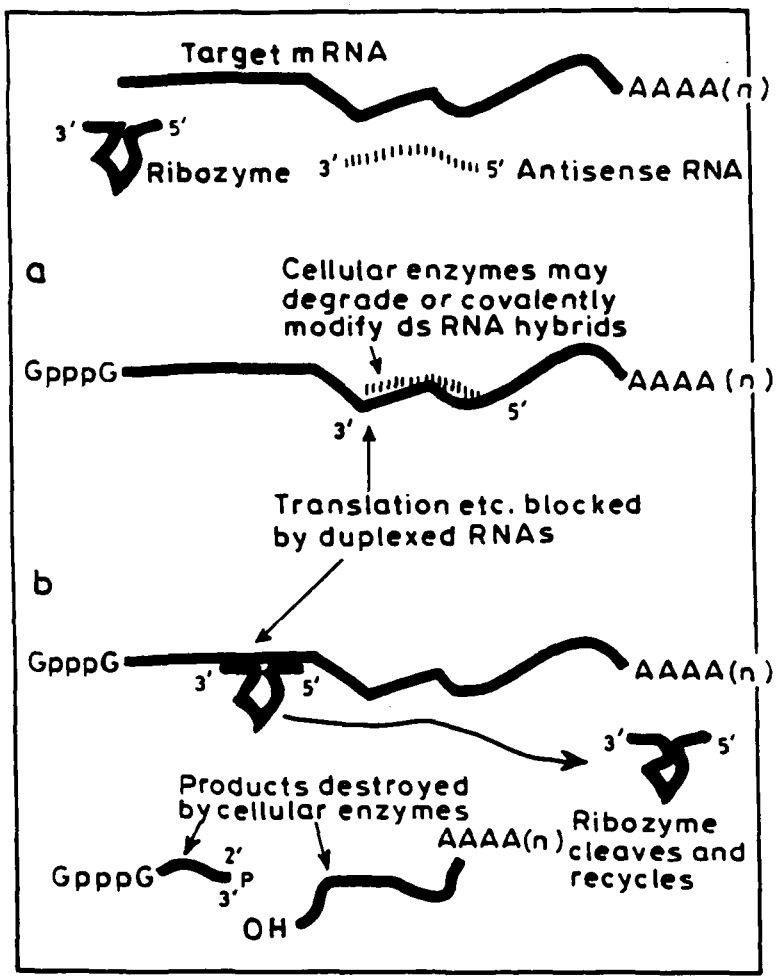


Figure 6: Ribozyme versus antisense-RNA-mediated inhibition of mRNA expression. The targeted messenger base pairs with either (a) the antisense RNA or (b) the ribozyme [Rossi, 1995]


Target mRNA
AAAA(n)
3'
5'
Ribozyme
3'
5' Antisense RNA
a
Cellular enzymes may degrade or covalently modify ds RNA hybrids
GpppG
AAAA (n)
3'
5'
Translation etc. blocked by duplexed RNAs
b
GpppG
3'
5'
AAAA(n)
3'
5'
Products destroyed by cellular enzymes
Ribozyme cleaves and recycles
AAAA(n)
GpppG
2'
P
3'
OH

CHAPTER II: General materials and methods

CHEMICALS

The following chemicals were used:

MATERIALS:	SOURCES
Acetic acid	SRL, India
Acrylamide	Amresco, USA
Agar agar	Hi-Media, India
Ammonia aqueous	Fluka, USA
Ammonium chloride	SRL, India
Ammonium persulphate	Serva Feinbiochemica, Finland
Bis-acrylamide	Amresco, USA
Boric acid	E.Merck, India
Bovine serum albumin	Sigma, USA
Bromophenol blue	Bio-Red, USA
Calcium chloride	SRL, India
Casein hydrolysate (Vitamin free)	Difco Laboratories, USA
Chloramphenicol	Ranbaxy, India
Chloroform	SRL, India
Cytidine	Sigma, USA
Diethylpyrocarbonate (DEPC)	Sigma, USA
Dithiothereitol (DTT)	Sigma, USA
DNase free pancreatic RNase	Promega, USA
Ethanol (Absolute alcohol)	E.Merck, India
Ethedium bromide	Sigma, USA
Ethylenediamine tetra-acetate	SRL, India
Ethylmethane sulphonate (EMS)	Sigma, USA

Ficol	Sigma, USA
Formaldehyde	Sarabhai,M.Chemical, India
Formamide	SRL, India
Glacial acetic acid	E.Merck, India
Glucose	Qualigen, India
Glycerol	SRL, India
β-Glycerophosphate	Sigma, USA
Hydrochloric acid	SRL, India
8-hydroxyquinoline	Sigma, USA
Isoamyl alcohol	SRL, India
Liquid nitrogen	ICGEB, India
Lysozyme	Sigma, USA
Magnesium acetate	SRL, India
Magnesium chloride	SRL, India
β-mercaptoethanol	Sigma, USA
Nutrient broth	Hi-media, India
Phenol	SRL, India
Polyvinylpyrrolidine	Sigma, USA
Potassium chloride	SRL, India
RNase free DNase I	Promega, USA
RNAsin	Promega, USA
Sodium acetate	SRL, India
Sodium chloride	Hi-Media, India
Sodium citrate	SRL, India
Sodium dodecyl sulphate (SDS)	SRL, India
Sodium hydroxide	SRL, India
Sodium phosphate	SRL, India

MEDIA

The composition of the GPTG medium is as under:

β-Glycerophosphate	10 g/l
NaCl	5.8 g/l
KCl	3.7 g/l
$CaCl_2$	0.11 g/l
$MgCl_2.6\ H_2O$	0.1 g/l
NH_4Cl	0.1 g/l
Trizma (tris) base (pH 7.4)	12.1 g/l
Na_2SO_4	0.14 g/l
Glucose (50%)	10 ml
Vitamin free Casein Hydrolysate (10%)	10 ml
d_2H_2O	980 ml

1. For GPTG plates 15 g of agar-agar was added, autoclaved and then 20 μg/ml of tryptophan and cytidine was added.

2. For top agar concentration of agar-agar was reduced to half (7.5%).

Nutrient Broth:

Peptone	5 g/l
NaCl	5 g/l
Beef extract	1.5 g/l
Yeast extract	1.5 g/l
pH (approx.)	7.4

Soft Agar:

Nutrient broth	13 g/l
Agar-agar	7 g/l

Hard Agar:

Nutrient broth	13 g/l
Agar-agar	14 g/l

Solutions for RNA extraction:

Solution I:

10mM Tris-HCl(pH 7.5) + 1mM EDTA	150 µl
1M Dithiothereitol	1.5 µl
20 U/µl RNasin	1.5 µl
10 mg/ml Lysozyme	4.0 µl

Solution II:

1M $MgAc_2$	20.0 µl
13 mg/ml RNase free DNase I	17.5 µl
20 U/µl RNasin	1.00 µl
DEPC-treated Water	61.5 µl

Solutions for oligonucleotide End-labeling:

Tris-EDTA buffer (TE):

Tris-HCl	10 mM
EDTA	1.0 mM
pH 7.5	

5X Kinase buffer (5X KB):

Dithiothereitol (DTT)	25 mM
$MgCl_2$	50 mM
Spermidine	1.0 mM
Tris-HCl	0.5 mM
pH 8.0	

Diethylpyrocarbonate-treated water:

DEPC	0.1% (v/v)

Solutions for hybridization:

Prehybridization mix:

Formamide	50%
SSPE	5X
Denhardt's reagent	2X
SDS	0.1%

SSC (20X):

NaCl	3.0 M
Na-Citrate	3.0 M
pH 7.5 adjusted with 10N NaOH	

Denhardt's solution (50X):

Ficol 400	5 g

Polyvinylpyrrolidine	5 g
BSA	5 g
DEPC Water	500 ml

SSPE (20X):

NaCl	175.3 g
NaH_2PO_4	24.6 g
EDTA	7.4 g
H_2O	800 ml

pH 7.4 was adjusted with NaOH, volume make to 1000 ml

Composition for extraction buffer:

NaCl	0.5 M
Tris	20 mM
EDTA	1 mM
pH (8.0)	

Solutions for DNA mini-preparation:

Solution I:

Glucose	50 mM
Tris-Cl	25 mM
EDTA	10 mM
pH (8.0)	

Solution II:

NaOH	0.2
SDS	1 %

Solution III:

Potassium acetate	5 M
Glacial acetic acid	11.5 ml
H_2O	88.5 ml

Phage and Bcterial Strains:

T4Do, an osmotic shock resistant mutant of T4D was used as standard (wild type) T4.

The *frd*1 mutant was isolated from T4Do strain after Hydroxylamine mutagenesis exploiting the white halo phenotype on the OK305 strain of *E.coli* B . A series of *frd*1*nrdB* double mutants was used for isolation of revertants .

Bacterial strains and plasmids:

E.coli B cells were used for phage crosses. *E. coli* BB cells were used for preparation of phage lysates (stock). Genetic selection and screening for T4 *td*, *frd*1 and *nrdB* mutations were routinely performed by plating suspected mutants on *E.coli* OK305 cells. OK305 is uracil-requiring *E.coli* B mutant that has very low cytidine deaminase activity and no detectible deoxycytidine deaminase activity. *E.coli* JM101 and JM109 strains are F^+ strain suitable as hosts for M13mp10 phage infection. JM109 Thy A is a thymine- requiring JM109 derivative isolated in the laboratory of Marlene Belfort, SUNY, Albany, NY.

The *nrdB* series of plasmid subclones used for marker rescue experiments were obtained by cloned fragments of the T4 *nrdB* gene into plasmid pBSM13^{+} (Stratagene) followed by transformation of the plasmid insert construct into JM101 and JM109. The *nrdB* subclone library was constructed in the laboratory of Dr. David Shub, SUNY, NY.

LIST OF RECOMBINANT CLONES

Plasmid	Size	Cloned region
pJSS10	4.6 Kb	nrdB gene
pJBK1	3.9 Kb	nrdB intron
pJG106	3.8 Kb	A+B region of intron
pJG108	3.6 Kb	C+D region of intron
pJSK7	3.3 Kb	D region of intron
pJSR1	3.5 Kb	B region of intron
pJBS1	3.4 Kb	A region of intron

METHODS

Maintenance and growth of bacterial strains:

Each strain of *E.coli* was streaked on a nutrient agar plate. A single colony was picked up from the plate and repurified again by streaking. The culture was tested on the basis of associated genetic markers raising it from a single colony on the master plate. Having satisfied with the test clone,

the culture was streaked on nutrient agar slants. It was then allowed to grow overnight at 37°C and stored at 4°C. The culture was transferred on to the fresh slants every month.

Stab preparation:

Stab cultures were also prepared from the cultures originated from the single colony of checked associated marker. These stabs were stored at 4°C for a year or so.

Gycerol culture:

Glycerol cultures were made by mixing 0.8 ml of freshly grown overnight cells into 0.2 ml of 15% sterilized glycerol. It was preserved at -70°C for several years.

PREPARATION AND PRESERVATION OF PHAGE STOCKS

Lysate preparation:

Over night Liquid Lysate: Over night liquid lysates were used for general phage stocks and usually yielded a titre of 1x10 to $1x10^{10}$ plaque forming units per ml (PFU/ml) . Over night liquid lysates were prepared by inoculating 5 ml of liquid broth medium with either a freshly picked plaque containing approximately 10^7 PFU or 10^7 phage particles from another stock. A 0.1 ml of fresh, full grown *E. coli* BB cells were added to the broth containing the phage and the tube were incubated at 37°C in a dry air incubator for 15-19 hours. The lysate were then chloroformed (12 drops) and allowed to sit at room temperature for 15 minutes before dilution and titre

determination. The lysate was stored at 4°C.

Plate Lysate: High titre phage stocks were made using plate lysates which yielded a titre of $1x10^{10}$ to $1x10^{12}$ PFU/ml. Approximately $2x10^5$ phage particles and $2\text{-}4x10^8$ *E. coli* BB cells were added to top agar and subsequently poured onto the LB agar plate. The LB medium was used for making plate lysates. Plates were incubated at 37°C for 12-14 hours at which point the plaques were partially confluent on the plate with only little of the cell lawn remaining visible. One ml of chloroform was spread over the top of the plate so that the entire surface of the plate was covered with chloroform. After the chloroform dried on the plate, 3ml of LB was poured over the top of the plate. The plate was briefly swirled at room temperature by hand every half an hour or so for three hours before the LB medium was collected from the top of plate with pasteur pipette and placed in a 15 ml corex centrifuge tube. In some cases two or three plates were made for each phage stock so that the final volume of high titre lysate approached to 10 mls. All plate lysates were centrifuged for 15 minutes at 7000 rpm in a Sorvell SS-34 rotor at 4°C. The supernatant was collected and few drops of chloroform were added. This lysate was stored at 4°C.

Bubbling Lysate Technique: Large (20 ml) bubbling lysates were prepared by using a large bubbling tube containing 20 ml of LB media, inoculated with 0.8 ml cells from a fresh over night culture of *E. coli* B or BB and aerated at 37°C until the cells were reached 1 to 2 X 10^8/ml. The culture was then inoculated with 2 to 4 X 10^6 phage particles from a stock lysate or with 1 ml of a 2 ml fresh plaque picking in diluting fluid. Aeration at 37°C continued until the lysate cleared, no longer than 5 hrs. At that point 1 ml of chloroform was added with aeration at 37°C continuing for an additional ten minutes. The lysate was transferred to a 45 ml polypropylene tube and

centrifuged at 7000 rpm in Sorvall SS-34 rotor for 30 min at 4°C to pellet down the cell debris. The supernatant was carefully decanted and saved. After addition of several drops of chloroform, the lysate was diluted, assayed and then stored at 4°C. The titre was typically between 10^{10} and 10^{11} plaque forming units per ml.

Mutagenesis:

*frd*1 mutant of T4 phage was further mutagenized by EMS. A 0.2 M solution of EMS was mixed with 12.5 % ethanol which was then incubated along with the equal volume of T4 *frd*1 mutant for 2 hours at 37°C followed by diluting it upto 10 volumes of nutrient broth, containing $Na_2S_2O_3$. This reaction mixture was further incubated for 24 hours at 37°C (Drake and Greening, 1970).

Phage plating conditions:

In general phages were visualized as plaques on agar plates by either one of the two methods mentioned below:

(i) Adding phage and *E.coli* cell to 2.5 ml of top agar prior to spreading the agar on an agar plate.

(ii) Spotting a small volume of phage (30 ml) on an agar plate seeded with the host bacteria. After plating or spotting of phage, plates were usually placed in a 37°C incubator for over night (10-14 hours) and the plaques were observed the next day.

td⁻/frd⁻ plating conditions:

T4 mutants defective in thymidylate synthase production (*td⁻*) or defective in dihydrofolate reductase production (*frd⁻*) were screened by assaying for a white halo plaque phenotype.

White halo plaque phenotype:

The white halo appears around a plaque, under special plating conditions with *E.coli* OK305 cells which have defect in their *de novo* cytidine deaminase activity and deoxycytidine deaminase activity. When a lawn of OK305 was grown on petri dishes with synthetic medium (GPTG) (Hall et al, 1967) containing cytidine as a sole pyrimidine source, there was thus a dUMP deficiency which could be remedied if the cells were infected with a mutant defective either in dTMP synthase (*td⁻*) or in dihydrofolate reductase (*frd⁻*). Both mutants cause dUMP accumulation which can leak from infected cells and feed neighbouring cells. This causes mutant plaques to be surrounded by a white halos.

Phage particles were plated with $2x10^8$ *E. coli* OK305 cells on GPTG+CR plates (GPTG plates with 20μg/ml CR added to the agar before plate pouring) supplemented with 4μl of CR (10 mg/ml) added to the top agar for each plate. After the top agar hardened, plates were incubated at 37°C for 8-10 hours. Plates can be incubated at 37°C for a short period of time. Generally, *frd*1 mutations result in a stronger white halo than do *td* mutation.

nrdB⁻ plating condition:

A second mutation in the *td⁻* or *frd1⁻* T4 phage in either of the *nrdB* gene , the cd gene or gene 56 will result in the disappearance of the halo which was present when there was only the initial *td⁻* or *frd1⁻* mutation. To isolate *nrdB⁻* mutant phage, $2x10^8$ OK305 cells ,40μg of CR along with *frd⁻/ td⁻* phage that had been mutagenized for a second mutation were added to GPTG top agar and then poured on to GPTG+CR agar plates. T4 *nrdB⁻*, *cd⁻* or gene 56 mutants carrying the initial *frd/td* mutation, did not exhibit the white halo phenotype under these conditions while *nrdB⁺* phage, wild type, *td⁻* and *frd1⁻* phage did produce a white halo.

Plasmid mini-preparation:

For mini preparation of plasmid DNA, a slightly modified procedure for rapid alkaline extraction method of Birnboim and Doly (1980) was followed.

Harvesting:

A single colony was transferred into 2 ml of LB medium containing the appropriate antibiotic in a loosely capped 15 ml tube. The culture was incubated overnight at 37°C with vigorous shaking. Then 1.5 ml of over night culture was poured into microfuge tube followed by centrifugation at 14000 rpm for 30 seconds at 4°C in microfuge. The medium was removed leaving the bacterial pellet as dry as possible.

Lysis by alkali:

Bacterial pellet was resuspended in 100μl of ice cold solution I (as described in general materials and methods) by vigorous vortexing. A 200μl of freshly prepared solution II was added followed by inversion of the tubes

rapidly five times (vortexing not allowed). Then 150µl of ice cold solution III was added and gentle vortexing was done for 10 seconds. The tubes were then stored on ice for 3-5 minutes followed by centrifugation at 14000 rpm for 5 min at 4°C in a microfuge. The supernatant were transferred to fresh tubes. An equal volume of phenol:chloroform was then added and it was mixed by vortexing and centrifuging at 14000 rpm for 2 min at 4°C in microfuge. The upper layer was transferred to fresh tube. DNA was precipitated in 2 volumes of ethanol at room temperature. It was then mixed by vortexing followed by incubation of the mixture at room temperature for 2 min. The mixture was centrifuged at 14000 rpm for 5 min at 4°C in a microfuge. The supernatant was discarded and the pellet of DNA was washed twice with 1 ml of 70% ethanol at 4°C. The pellet thus obtained was air dried, and dissolved in TE (pH 8.0).

Transformation of *E.coli* cells with pJSS10, pJBK1, pJG106, pJG108, pJSR1, pJBS1 and pJSK7 plasmid DNA:

The transformation of these plasmids carrying clones of *nrdB* gene and its various intronic regions was carried out into the JM101 cells.

Preparation of competent cells:

A single colony of JM101 strain was picked from a freshly grown plate and inoculated into 50 ml of Luria broth in a 250 ml flask. The cells were grown with vigorous shaking at 37°C for about 3 hours and growth was monitored at 600 nm (density should not exceed 10^8cells/ml). These cells were then transferred to a pre-chilled sterile 50 ml polypropylene tube and stored on ice for about 10 min followed by centrifugation at 4000 rpm for 10 min at 4°C The cell pellet was thoroughly washed once with chilled 50 mM $CaCl_2$ and resuspended again in 30 ml of the $CaCl_2$ solution. The cells were kept

on ice for 1 hour followed by centrifugation in cold and resuspended in 1 ml of the ice-cold $CaCl_2$ solution.

Transformation of competent cells:

20µl plasmid and 200µl competent cells were added with 100µl of 100mM tris-Cl, pH 7.5, which was then incubated on ice for 30 min Heat shock was given to cells 42°C for 2 min. It was chilled on ice for 5 min, centrifuged at 4000 rpm for 5 min at room temperature and supernatant was discarded. Cells were resuspended in 1 ml of fresh LB and incubated at 37°C for 1 hr. Again it was centrifuged at 4000 rpm for 2 min. 100µl of fresh LB was added, resuspended and spread on a LB plate containing antibiotic. Grown overnight at 37°C.

Marker rescue mapping procedure:

Marker rescue:

The recombination based marker rescue technique was used to map *nrdB* mutations. Marker rescue was said to be positive if a mutant phage was infected into a cell which had a corresponding wild sequence for that mutation in the phage genome. Some percent of the progeny phage would have picked up this wild type sequence from the cell and thus showed wild type phenotype. Using this principle, the *nrdB*$^-$ (non-halo) mutants were identified by infecting host cells containing a cloned *nrdB* gene and looking for *nrdB*$^+$ (halo$^+$) progeny in the marker rescue lysate. Identification of intron mutants and their finer mapping was also conducted using the same principle of marker rescue by using a series of intron subclones of the *nrdB* gene.

Isolation and purification of EMS-induced mutants:

EMS mutagenized *frd*1 mutants of T4 phage were plated under the GPTG condition with OK305 cells. Halo$^-$ plaque was picked and suspended in 1 ml of diluting fluid. This isolated plaque was tested for purity and hence those plaques were picked that did not show any halo$^+$ character even if 10,000 plaques were screened. A purified well isolated halo$^-$ plaque was preserved in diluting fluid with a few drops of chloroform and this lysate was used to prepare a high titre lysate for further studies.

Isolation of *nrdB* mutants:

EMS mutagenized *frd*1 mutants of T4 phage was infected on JM101 cell containing the clone of nrdB gene. The cells with phage were plated on to the nutrient agar plate. After the incubation of several hours at 37°C, plaques were appeared. Two well isolated plaques were picked and each dissolved in 1 ml of diluting fluid. A few drops of chloroform were added and preserved at 4°C. This isolate was then tested for halo phenotype under the GPTG plating conditions with OK305 cells. The strategy was that, if it showed halo$^+$ phenotype (>5%) then the mutant stock was said to be *nrdB* mutant.

Mapping of *nrdB* intron mutants:

Phage stocks from the *nrdB* mutants were diluted appropriately and plated on nutrient agar with JM101 cells containing the pJBK1 clone carrying the *nrdB* intron. For each one of the marker rescue plates, two well isolated plaques were picked into 1 ml diluting fluid and three drops of chloroform were added to kill the cells. Serial dilutions were made and tested for halo$^+$ phenotype. Appropriate controls with *frd*1 mutant and wt phage

were plated for comparison with marker rescue plates. A minimum of 10^4 plaques were screened for each marker rescue. The recombination based marker rescue technique was said to be positive if a mutant phage was infected into a cell which had a corresponding wt sequence for that mutation in the phage genome. Some percent of progeny phage would have picked up this wt sequence by recombination with the insert in the plasmid and thus showed wt (halo$^+$) phenotype.

Mapping of sub-regions of the *nrdB* intron mutants:

Phage stocks from the *nrdB* mutants were diluted appropriately and plated on nutrient agar with JM101 cells containing the pJG106 carrying A+B region of intron. For each marker rescue plates two well isolated plaques were picked into 1 ml diluting fluid and three drops of chloroform were added to kill the cells. Serial dilutions were made and tested for halo$^+$ phenotype. Appropriate control with *frd*1 mutant and wt phage were plated for comparison with marker rescue plates. A minimum of 10^4 plaques were screened for each marker rescue. The recombination based marker rescue technique was said to be positive if a mutant phage was infected into a cell which had a corresponding wt sequence for that mutation in the phage genome. Some percent of progeny phage would have picked up this wt sequence by recombination with the insert in the plasmid and thus showed wt (halo$^+$) phenotype. Similar experiments were performed for the mapping of remaining region of intron with different subclones of *nrdB* intron. The pJG108 was used for mapping the mutants within C+D region, pJBS1, pJSR1, and pJSK7 were used for the mapping of A,B, and D region of intron respectivey.

RNA preparation:

E. coli BB cells from a freshly grown culture were diluted 100-fold into Luria broth and the culture was allowed to grow again at 30°C in a shaking water bath until the cell density reached to $2x10^8$ cells/ml. The cells were transferred to 50 ml Oak Ridge centrifuge tube and centrifuged at 10,000 rpm for 10 min at 4°C. The cells were resuspended in 0.5 ml LB followed by the addition of the infecting phage and placed the tubes on ice for 5 min. Cells were usually infected with a phage multiplicity of infection (MOI) of at least seven. The cell/phage mixture was added to 30 ml of pre-warmed (30°) LB and incubated at 30°C for 9 minutes in shaking water bath. 1.2 ml of chloramphenicol was added at 2.5mg/ml and placed the cells at 30° for one minute. The flask containing the culture was then rapidly cooled in an ice-water bath until the temperature of the culture had reached 4°C. The cells were then transferred to pre-chilled centrifuge tubes and spun at 10,000 rpm for 10 min at 4°C followed by resuspension in one ml of 10 mM tris (pH 7.5)-CAM (100μg/ml) solution without vortexing. These were then transferred to 1.5 ml eppendorf microfuge tube and centrifuged at 4°C for 2 min. Supernatant was discarded and resuspended the cells pellet by vortexing in 0.5 ml of the Tris-CAM solution. It was again centrifuged for 2 min and supernatant was discarded. The pellet was blotted with tissue paper, followed by freezing the pellet in liquid nitrogen. It was used immediately or stored at -80°C for upto a month.

RNA Extraction:

A 157μl of solution I (as described in the methods section) was added to the frozen pellets and resuspended (vortexing allowed at this stage). The suspension was quickly frozen and thawed three times in liquid

nitrogen and then 20µl of solution II was added . The mixture was shaken gently and placed on ice for 45 min. Then 10µl of 0.2% acetic acid and 10µl of 10% SDS was added. The tubes were placed at room temperature for 5 min. Three series of organic extractions were usually performed with phenol / chloroform / isoamyl alcohol in a 25:24:1 ratio using phenol, buffered with 0.1 M tris-HCl pH 7.5, 0.1% 8-hydroxyquinoline and 0.2% mercaptoethanol . The tubes were centrifuged for 3 min at 14000 rpm at 4°C. The aqueous (top) phase was saved into fresh eppendorf tube and used for the second extraction. The second extraction was carried out by chloroform / isoamyl alcohol in 24:1 ratio. Again the top phase was saved (after centrifugation as above) for the third extraction. After third extraction, approximately 100µl was transferred to a new 1.5 ml tube and 1/100th volume of 1M $MgAc_2$, 1/10th volume of 3M NaAc (pH 5.2) and 2.5 ml of 100% ethanol was added. Precipitation of RNA was allowed to occur at -80°C for at least one hour or -20°C for over night. Tubes were then centrifuged at 4°C for 20 minutes. The pellet was then resuspended in 100µl of DEPC-treated water followed by the addition of 1µl of 1M $MgAc_2$,10µl of 3M NaAc (pH 5.2) and 2.5 ml of 100% ethanol. Again it was precipitated at -20°C or -80°C followed by centrifugation. The pellet was washed with 70% ethanol and dried in vacuum. The pellet was resuspended in appropriate volume of DEPC-water. RNA concentration was measured at 260 nm by making dilution. A concentration factor of 1 OD_{260} = 40µg/ml RNA.

End-labeling of synthetic oligonucleotides:

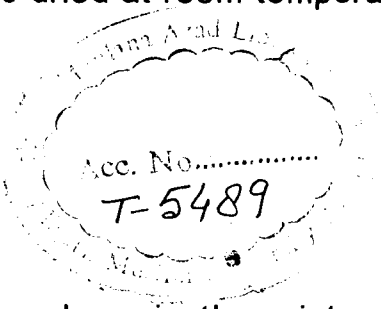
5'-End labelling of oligonucleotides with γ^{32}P-ATP was performed. Twenty picomoles of purified synthetic oligonucleotide , 1.0µl of T4 polynucleotide kinase at (10 U/µl), 2.0µl of 5X Kinase buffer, 0.8µl of DEPC-treated water and 4.2µl of γ^{32}P-ATP(3000 ci/mM) were added to a 1.5ml

tube. The kinase mixture was incubated a 30°C for 45 minutes followed by incubation at 95°C for 3 min. After cooling on ice, 25µl of DEPC-treated water was added.

Urea-PAGE obtained by 8% acrylamide / bis-acrylamide and 7M urea was run for 15-20 minutes followed by the exposure on X-ray film, the band was cut from the gel and oligo was eluted in the extraction buffer. The solution was filtered via siliconized glass wool.

Dot blot hybridization:

A piece of Hybond membrane was completely soaked in water for 10min and then soaked in 20X SSC solution for one hour at room temperature. Mean while the manifold hybridization apparatus was cleaned with 0.1 N NaOH and then rinsed it well with sterile water. All the slots were filled with 10X SSC, and a gentle suction was applied until all the fluid has passed through the membrane filter. The sample was prepared by mixing the RNA with 20µl of 100% formamide, 7µl of 37% formaldehyde ,2µl of 20X SSC, the mixture was then incubated for 15 minutes at 68°C and then it was cooled on ice. It was then loaded into separate slots and gentle suction was applied . Each slot was rinsed with 1 ml of 10X SSC after all the samples have passed through slots. The membrane was dried at room temperature, and crosslinked under the UV light.



Hybridization and autoradiography:

The filter was prehybridized for one hour in the mixture of 50% formamide, 5X SSPE ,2X Denhardt's reagent and 0.1% SDS at 42°C. Denatured radiolabeled probe was added directly to prehybridization fluid. To

detect the low abundance mRNA, 0.1μg of probe was used, whose specific activity exceeded $2x10^8$ cpm/μg . Incubation was continued for 16-24 hours at 42°C. Then the filter was washed for 20 min at room temperature in 1X SSC with 0.1% SDS followed by three washes of 20 min each at 68° C in 0.2X SSC with 0.1% SDS .

An autoradiogram was obtained by exposing the membrane for appropriate period of time.

CHAPTER III: Isolation and mapping of EMS induced splicing defective point mutations of *nrdB* intron of bacteriophage T4

INTRODUCTION

Bacteriophage T4 contains three self splicing introns, the *td* intron (chu et al, 1984), the *nrdB* intron (Sjoberg et al, 1986) and *sunY*(*nrdD*) intron (Gott et al, 1986). These introns differ substantially both in sequence and in size. The *nrdD* and *td* introns respectively comprise 1033 and 1016 nucleotides while the *nrdB* intron is only 598 nucleotides long.

The white halo plaque phenotype exhibited by *frd*1 mutants under special plating conditions (Hall and Tessman, 1967) is an excellant system for the isolation of *nrdB* mutants which can be characterized further for intron mutants by using the *nrdB* intron subclones to rescue the intron mutation.

We have mutagenized *frd*1 mutants by *in vitro* ethylmethane sulphonate treatment (Drake and Greening, 1970). Using this approach a number of *frd*1*nrdB* double mutants were isolated. In the previous study Lal and Hall (1993) used hydroxylamine as a mutagenic agent which is specific to cytosine, whereas EMS is non-specific alkylating agent that causes bidirectional transition i.e. AT to GC as well as GC to AT. EMS reacts with guanine, adenine, cytosine and 5-hydroxymethylcytosine when the free bases or nucleotides are exposed in solution (Brookes and Lawley, 1960, 1962). The reaction of this agent with DNA produces 7-ethylguanine and 3-ethyladenine and other products may occur to a lesser extent (Brookes and Lawley, 1962).

The pJSS10 clone carrying whole *nrdB* gene was used to isolate the *nrdB* mutants. The intron sequence along with 80 bases from 3'-end of the exon I and 4 bases from 5'-end of the exon II, cloned in pJBK1 was used to map the *nrdB* mutants within the short stretch of intron. Region A+B of the

intron was mapped by using subclone pJG106. The remaining intron mutants were mapped within C+D region by using the subclone pJG108. These two regions of intron were splitted into A,B,D, and C regions. The mutations were mapped within these regions by using subclones, pJBS1, pJSR1, and pJBK1 respectively. Region C does not have any clone, therefore, mutations in this region were mapped by default.

Intron of *nrdB* gene has a series of short conserved sequence elements P,Q,R and S. Interestingly, all the mutants isolated in this study clustered into regions of the *nrdB* intron containing at least a portion of the P,Q,R,and S sequences. The frequency of homologous recombination (crossing over occurs during the recombination of T4 mutant gene and wild type clone of the same gene) resulting in wild type phenotype has been used to predict the relative distances between various mutations. Dot blot splicing assay was also performed.

METHODS

The *frd1* mutants of phage T4 was mutagenised by ethylmethane sulphonate (Drake and Greening, 1970).

Isolation and mapping of *nrdB* intron mutants were performed on the basis of recombination based marker rescue technique (as explained in chapter II). In this chapter we have mutagenized the *frd*1 mutants which have halo$^+$ phenotype as described in the section of general methods. These *frd*1 double mutants were then screened for purification and isolation of the *nrdB* mutants. The mapping was carried out by using recombinant plasmids containing the desired regions of *nrdB* intron and their sub-regions to be mapped. These recombinant plasmids were transformed into JM101 cells.

Subsequently the mutagenized *frd*1 double mutants were diluted appropriately and plated on LB plates with the recombinant clones (Fig. 6). For each one of the marker rescue plates, two well isolated plaques were picked into 1ml of the diluting fluid and three drops of chloroform were added to kill the cells, several dilutions were made and tested for halo$^+$ phenotype. Halo is excess growth of cells around the plaques due to accumulation of uracil in *frd*1 mutants of T4 phage as described in general methods (Fig. 2). Appropriate control plates with *frd*1$^-$ and wt phage were plated for comparison along with the test marker rescue plates. A minimum of 10^4 plaques were screened for each marker rescue. The recombination based marker rescue is said to be positive if a mutant phage infected a cell which had a corresponding wt sequence for that mutation in the phage genome. A fraction of the progeny phage would have picked up this wt sequence by homologous recombination from the insert in the plasmid and thus showed wt (halo$^+$) phenotype. The mutant which does not show wt (halo$^+$) phenotype with recombinant plasmid (not less than 5% halo$^+$ frequency) does not have any mutation within the region which have been cloned.

The total RNA extraction was initiated after a 9 min infection to maximize the amount of *nrdB* mRNA isolated from the selected *frd*1*nrdB* T4 mutant infected *E.coli* BB cells was performed as described by Lal and Hall (1997). RNA Dot blot hybridizations were also performed by using the protocol described by Belfort et al (1985).

RESULTS

Determination of mutagen dosage: The mutagenesis experiment was performed to check the efficacy of the EMS and, therefore, various concentrations of EMS were taken to determine the appropriate dose of EMS which would give the survival and mutagenicity within the range of estimation. The phage *frd*1 mutant obtained from Prof. D.H.Hall, (Gorgia Institute of Technology, Atlanta) was mutagenized by the EMS. After mutagenesis the phenotype of some of the *frd*1 mutants had changed from halo$^+$ to halo$^-$ (Fig. 1). The halo$^-$ phenotype is because of the defect in *nrdB* or *cd* gene that code for a small subunit of ribonucleotide reductase and cytidine deaminase respectively. Due to the inactivation of either of the gene synthesis of uracil is blocked and uracil will no more be accumulated in T4 phage which in turn would not release out of lysed phage particles and hence plaques are devoid of halo$^+$ phenotype (Fig. 2,3,4). The percent inactivation and mutational frequency of the EMS-treated *frd*1$^-$ mutants were found to vary with respect to the concentration of EMS. It also varied in the presence and absence of sodium thiosulphate which was used to quench the reaction (Table 1,2).

Isolation and purification of mutant plaques: The EMS treated *frd*1 mutants were plated on GPTG plates along with OK305 cells.Screening for appearance of halo$^-$ phenotype of the plaques was done. Approximately 800 halo$^-$ plaques were picked after EMS mutagenesis. Each plaque was suspended in 1ml of diluting fluid. The titre of all the mutants were checked and recorded as 10^6 to 10^7 PFU/ml. Appropriate dilutions of the mutant phage were then made and plated again on GPTG medium to check the purity of the isolated phage mutants. The isolated mutant was considered

to be pure if the plating of test mutant on GPTG did not result in the appearance of any halo$^+$ plaque out of 10^4 plaques screened for this purpose. A total of 432 *frd*1*nrdB* double (halo$^-$) mutants were thus purified.

Isolation of *nrdB* mutants: The *frd*1 double mutants (halo$^-$) obtained by EMS mutagenesis were further screened for the presence of a mutation within the *nrdB* gene (Fig. 5). The basis of selection of *nrdB* mutation was the marker rescue with the pJSS10 clone (Table 3). This clone contained the *nrdB* gene in the plasmid. This screening resulted in the isolation of 156 out of 432 test mutants within the *nrdB* gene (Table 4).

Mapping of *nrdB* intron mutants: The *nrdB* gene mutants thus obtained were mapped within the *nrdB* intron using the subclones of the *nrdB* gene . Genetic mapping of *nrdB* intron mutants were performed on the basis of the recombination based marker rescue technique (Fig. 5). A total of 31 mutations were mapped within or near the intron of *nrdB* gene, (Table 5,6).

The frequencies of halo$^+$ (wt) phenotype of these mutations with the pJBK1 (intron) subclone (Fig. 6) were recorded to be significant (Table 5). Since each mutation can be placed at its appropriate position within the intron on the basis of their frequencies, the lower recombination frequency of mutant with wt should always lie towards the end of the sequences whereas mutation at the centre would always show higher frequencies (Fig. 10). Most of the mutations showing lower frequencies were mapped near the ends of intron while 31 of the phage showed remarkably high frequencies of halo$^+$ and are mapped in the middle of the intron.

Mapping within JG106 (A+B) region of intron: The strategy for narrowing down the mapping of mutations within the subregions of intron was based

on the same principle that a sub-region of intron, cloned within the plasmid pJG106 (region;A+B) (Fig. 6) was used to rescue the wt character (halo$^+$) (Fig. 2) . Seven mutations showed positive marker rescue with the subclone pJG106 (Table 7). This strongly suggests that these seven mutations were located within the A+B region of the 5' end of *nrdB* intron (Table 8).

Mapping within JG108 (C+D) region of intron: All of the 31 mutants were checked with pJG108 (Fig. 5,6) subclone for the positive marker rescue (Table 9) . Out of 31 *frd*$^-$1*nrdB*$^-$ mutants, 24 mutants showed positive marker rescue with the pJG108 and hence were mapped within C + D region of the *nrdB* intron (Table 10).

Mapping within sub-regions of JG106 (A+B): The JG106 region (A+B) was further subdivided into two regions A and B for which subclone pJBS1 and pJSR1 were used for mapping by marker rescue (Fig. 5)

Mapping within JBS1 (A) region of intron: All of these seven mutants (Table 8) were checked for marker rescue with the subclone pJBS1 (Fig.5,6), All of them were showing positive marker rescue with pJBS1 (Table 11) suggesting that all of them were carrying mutations within or near the 5' end of the region A of *nrdB* intron (Table 12).

Mapping within JSR1 (B) region of intron: The second attempt of marker rescuing of all seven mutants, which have already mapped within JBS1 region, was carried out with the pJSR1 subclone (Fig. 6). None of them exhibited a positive marker rescue with this clone (Table 13) therefore, all of the seven mutations could be mapped within the A region of the *nrdB* intron.

on the same principle that a sub-region of intron, cloned within the plasmid pJG106 (region;A+B) (Fig. 6) was used to rescue the wt character (halo$^+$) (Fig. 2) . Seven mutations showed positive marker rescue with the subclone pJG106 (Table 7). This strongly suggests that these seven mutations were located within the A+B region of the 5' end of *nrdB* intron (Table 8).

Mapping within JG108 (C+D) region of intron: All of the 31 mutants were checked with pJG108 (Fig. 5,6) subclone for the positive marker rescue (Table 9) . Out of 31 *frd*$^-$1*nrdB*$^-$ mutants, 24 mutants showed positive marker rescue with the pJG108 and hence were mapped within C + D region of the *nrdB* intron (Table 10).

Mapping within sub-regions of JG106 (A+B): The JG106 region (A+B) was further subdivided into two regions A and B for which subclone pJBS1 and pJSR1 were used for mapping by marker rescue (Fig. 5)

Mapping within JBS1 (A) region of intron: All of these seven mutants (Table 8) were checked for marker rescue with the subclone pJBS1 (Fig.5,6), All of them were showing positive marker rescue with pJBS1 (Table 11) suggesting that all of them were carrying mutations within or near the 5' end of the region A of *nrdB* intron (Table 12).

Mapping within JSR1 (B) region of intron: The second attempt of marker rescuing of all seven mutants, which have already mapped within JBS1 region, was carried out with the pJSR1 subclone (Fig. 6). None of them exhibited a positive marker rescue with this clone (Table 13) therefore, all of the seven mutations could be mapped within the A region of the *nrdB* intron.

Mapping within sub-regions of pJG108 (C+D): The subclone pJG108, having C+D region cloned in it was further subdivided into two regions, C and D.

Mapping within JSK7 (D) region of intron: The fine structure mapping of these twenty four mutants was carried out with pJSK7 (Table 14) and it was determined that out of 24 mutants 13 were mapped within JSK7 region (Table 15). Since all these thirteen mutants showed positive marker rescue with pJSK7, these lie within the 3'end of the D region of the *nrdB* intron.

The remaining, eleven (24-13=11) mutants would map within the C region of *nrdB* intron by default (Table 16).

Dot blot hybridization splicing assay : This hybridization assay was performed to analyse the splicing defective point mutations within the *nrdB* intron of bacteriophage T4. End-labelled DNA probes were designed such that their bindings were dependent on the splicing of the *nrdB* intron. Four DNA probes were used in this splicing assay (Table 17). Appropriate wild type T4 RNA preparation was used to measure splicing proficiency.

The ExII probe of 22 nucleotides was designed to anneal the nucleotide bases from 1865 to 1886 of Exon II region. It was used to ensure the presence of *nrdB* mRNA in the preparations. The second probe, ExI-ExII was used to bind at the splice junction specifically to the sequence starting from 1189 to 1200 of exon I and 1799 to 1811 of exon II. The third probe, ExI-IVS was designed to bind with the 5'-cutting site at sequence starting from 1189 to 1211 nucleotide bases of exon I and intron. The fourth probe, IVS-ExII on the other hand was designed to bind with the 3'-cutting site at the sequence starting from 1788 to 1810 nucleotide bases of intron and exon II. The ExII

probe served as a positive control because it will bind to all spliced and unspliced *nrdB* RNA. None of the RNA preparations displayed negative response with this probe (Fig. 7,8,9). Positive binding of the probe ExI-ExII would show the excision of intron and ligation of exons. And negative binding (i.e. no binding) would represent any defect in splicing. Positive binding of the probe, ExI-IVS would imply the defect at the 5'-cutting site of exon I and intron whereas positive binding of IVS-ExII probe would show splicing defect at the 3'-cutting site of intron and exon II.

Dot blot hybridization splicing assay of mutations in the region A of the *nrdB* intron: All the 7 mutants in the region A except AM668 exhibited negative binding with the probe ExI-ExII, while the AM668 showed positive binding with this probe similar to wild type (Fig. 7). The probe ExI-IVS displayed positive binding with all the mutants of this region, while wt exhibited negative binding as shown in fig. 7. The fourth probe, IVS-ExII did not show any binding to any of the mutants of this region (Fig. 7). This dot blot assay suggests that all the 7 mutants are splicing defective at their 5'cutting site but AM668 seems to be partially splicing defective at the 5'-cutting site (Table 18).

Dot blot splicing assay of mutants in region C of the *nrdB* intron: Out of 11 mutants, AM152 and AM779 showed positive binding with the probe ExI-ExII, and the rest exhibited negative binding (Fig. 8). All the mutants of the region C showed positive binding with the probe ExI-IVS. The mutants, AM11, AM753, AM764 and AM779 displayed positive binding with the probe IVS-ExII whereas AM38, AM49, AM166, AM326, AM429, AM152, and AM775 did not show any binding with this probe (Fig. 8). One mutant, AM152 was found to be partially defective in splicing at the 5'cutting site whereas AM779 seems to have partial defect at both the 5' and 3' cutting sites. The mutant

AM11, AM753, AM764 would be splicing defective at both the cutting sites, and the remaining mutants have a defect at their 5' cutting site (Table 19).

Dot blot splicing assay of mutants in the D region of the *nrdB* intron: There were 13 mutants in region D, (Table 20) out of these only three mutants, AM442, AM717, and AM730 displayed a positive response with the probe, EXI-EXII. The remaining mutants showed negative binding with this probe. The probe ExI-IVS resulted in positive binding with the mutants AM112 and AM153 (Fig. 9). As far as the probe IVS-EXII was concerned, all the mutants except AM153 displayed a positive response with this probe (Fig. 9). The mutant AM112 was splicing defective at both the 5' and 3' cutting sites whereas AM153 was found to be defective at its 5'cutting site only. Moreover, the mutants AM442, AM717, AM730 seems to have partial defects at the 3'cutting site. The remaining (AM342, AM425, AM445, AM484, AM487, AM492, AM719, AM724) were defective at the 3'splice site (Table 20).

Measurement of the genetic distances between the *nrdB* intron mutations: The frequencies of the halo$^+$ phenotype recombinants were used to calculate the relative genetic distances between mutants that had point mutations mapping within the same marker rescue clone/subclone as shown in fig. 6. For every genetic cross, approximately 5000 non-halo plaques were screened and the total number of halo$^+$ plaques for each mutant showing positive marker rescue were used to predict the genetic distance (Fig. 10) of that mutation from the end of the clone / subclone and also the relative distances of mutants with each other (Table 21,22,23).

DISCUSSION

As in the protein enzymes, the folding of the intron results in the formation of an active site juxtaposing key residues that are widely separated in primary sequence. This RNA structure catalyses splicing by bringing the 5' and 3' splice sites and guanosine into close proximity, and by activating the phosphodiester bonds at the splice site (Cech, 1990). Different group I introns, including the *nrdB* intron, have relatively little sequence similarity, but all share a series of the short conserved sequence elements P,Q,R and S with part of P/Q and R/S base pairing in conserved structure (Cech, 1988). These conserved elements are known to be important, because they are sites of cis-acting splicing mutations. This is not to say that the less or non-conserved structure will be functionally irrelevent, they could serve to position the catalytic residues or to stabilize the overall structure of the intron, or they could provide binding site for proteins that facilitate or regulate RNA splicing *in vivo* (Waring and Davies, 1984; Cech, 1988; Lal and Hall, 1997). This sequence complementarity between pairs of elements may allow them to interact to form a complex higher order structure. This complex structure appears to be essential for splicing activity. The P and Q sequences are closer to the 5' splice site of the intron whereas R and S sequences are closer to the 3' splice site of the intron. The region A contains both P and Q sequences whereas region B does not contain any of the conserved sequences. About half of the R sequence that is less conserved than the region A is carried by region C . Region D encompasses the remaining half of the R sequence and the whole of S sequence (Lal and Hall, 1993).

In the present study we have mutagenized *frd*1 T4 mutants with ethylmethane sulphonate. Previously, several workers have used

hydroxylamine and nitrous acid for the mutagenesis of *nrdB, frd*1 and *td* mutants (Hall et al, 1987; Lal and Hall, 1993; Kwon et al, 1995). Hydroxylamine is specific for cytosine and hence it induces unidirectional mutagenesis whereas EMS induces bidirectional GC to AT and AT to GC transitions. The GC to AT mutagenesis occurs at a remarkably higher rate than the AT to GC transition (Freese, 1961). We have isolated 31 EMS induced *nrdB* mutants as compared with 30 hydroxylamine induced mutants within the *nrdB* intron of T4 phage (Lal and Hall, 1993). The present study revealed that none of mutations were present within the B region of *nrdB* intron which was outside the conserved P, Q, R, and S sequences and hence our data suggest that only the conserved regions of the intron are involved in pre-mRNA splicing.

In our study 7 mutations were mapped in the region A similar to what was obtained by Lal and Hall (1993). However, unlike their study, we could not obtain any border line mutation in the region B. That mutation obtained by Lal and Hall (1993) was probably a border line mutation since it showed positive marker rescue with subclones corresponding to regions A as well as B, at a frequency of less than 5%. Cutting the *Sty*I site for subcloning produces staggered ends and thus a mutation in this site would show positive marker rescue with both subclones at a very low frequency. In addition, the region A contains both P and Q conserved sequences and the last base of the conserved sequence is the last base of subclone too. Thus there is strong possibility that the mutation is in the *Sty*I site. None of the mutations mapped in the B region as shown by Lal and Hall (1993). Eleven mutations mapped in the region C which is supposed to have less conserved sequences than that of A. Surprisingly, 3 out of 11 mutants within the region C displayed the 3'splice defect by dot blot assay (Table 19). Moreover the remarkable marker rescue frequencies with pJG108 clone carrying C+D

region of these mutants also suggest that the mutations lie within the non-conserved ORF region of the *nrdB* intron (Table 9). Lal and Hall (1997) did not find any mutation using hydroxylamine within the ORF region. This observation is consistent with the idea that the ORF region of *nrdB* intron could also play an important role in the splicing reaction. The absence of the conserved region might imply that tertiary interactions between the ORF and neighbouring conserved regions probably stabilize the RNA structure essential for splicing. Such an idea has also been put forward by previous workers in case of other group I introns (Inoue et al, 1985). These mutations within the suspected ORF region would have escaped the screening by Lal and Hall (1993) owing to limited target of hydroxylamine compared with the EMS targets. The hydroxylamine specifically attacks upon the C residues of the DNA thus resulting in GC to AT transition whereas EMS can attack upon all the DNA bases with greater or less intensity having more preference undoubtedly causing the GC to AT transition and also causing reverse transition of AT to GC. This also confirms to the total number of mutants within the C region to be 11, contrary to a total of 7 mutations within this region reported previously (Lal and Hall, 1993). These additional three mutants could have been the result of AT to GC transition. Interestingly the remaining 8 mutations in our case have mapped within the same small stretch (96 bases) of conserved nucleotides within the C region. Lal and Hall (1993, 1997) also found mutations within this stretch. The dot blot splicing assay and marker rescue frequency data support our contention (Table 19,22). Region D which contains a roughly equal amount of conserved sequences to region A has 13 mutations mapped in it.

These mapping results suggest that the conserved sequences in the *nrdB* intron, like the *td* and other group I introns, are important in pre-mRNA splicing (Michel et al, 1982; Davies et al, 1982; Cech et al, 1983).

Most of the mutations were mapped towards the 3'end of the *nrdB* intron compared to the *td* intron which has shown 15 mutations within 167 nucleotides of the 5'end and at least 12 mapped within 219 nucleotides of the 3'end of the intron (Hall et al, 1987). The present study also supports the findings of others that these are the ends of intron (5' and 3') which play a vital role in splicing by folding the secondary structure into a very precise tertiary structure configurations involving the complex P9.0 and P10 pairings (Hall et al, 1987; Lal and Hall, 1993, 1997).

Like many eukaryotic group I introns, the *nrdB* intervening sequence contains an ORF (Chu et al, 1986). The *nrdB* intron ORF (294 bases) is smaller than the *td* ORF (778 bases) which is again smaller than *nrdD* (788 bases) (Gott et al, 1988; Shub et al, 1988). The ORF sequences of other group I introns are not usually highly conserved and are looped out of secondary structure model for RNA folding (Waring and Davies, 1984). Similarly for the *td* intron these genetic studies imply that the intron ORF is not required for formation of the active splicing conformation (Hall et al, 1987). However, our findings reveal that the 3'end of the *nrdB* intron ORF is essential for splicing.

Seven mutations were mapped within the A region similar to that of Lal and Hall (1993). The sequencing analysis of those mutants clearly suggests the mutation sites of three mutants which were lying within exon I of *nrdB* gene. Interestingly the dot blot assay and MR frequency data favour the presence of at least three mutations existing very close to the 5'end of the intron in our case also. It is noteworthy that our clone pJBK1 for marker rescue assay was carrying 80 additional nucleotides of the Exon I.

Dot-blot splicing assay results suggest that most of the splicing defective mutants showed their defects at either 5' or 3' cutting site. While some of them were defective at both the cutting sites. Although some of the splicing defective mutants were mapped at the 3' end of intron yet they disrupt the 5' splicing activity. For instance mutant AM153 (Table 20) lies within the region D of the intron, inspite of that it showed splicing defect at 5' cutting site. It clearly suggests the folding of secondary structure of intron during splicing.

Table 1 : Effect of EMS treatment on the infectivity of *frd*1 mutants of T4 phage under extracellular conditions.

Concentration of mutagenic treatment to *frd*1 mutant of T4	No. of plaques on each plate					Titre (PFU/ml)	% Inactivation
	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-6}		
0.04M EMS + $Na_2S_2O_3$	-	-	2000	175	-	1.87 x10^7x 10 1.87x10^8	92
0.04M EMS - $Na_2S_2O_3$	-	-	-	565	4	4.8 x 10^7x10 4.8 x 10^8	79
0.1 M EMS + $Na_2S_2O_3$	-	3590	355	-	-	3.57 x 10^6x 10 3.57 x 10^7	98.43
0.1M EMS - $Na_2S_2O_3$	-	-	1040	220	-	1.62 x10^7x 10 1.62 x 10^8	93
0.2M EMS + $Na_2S_2O_3$	2000	170	-	-	-	1.85 x10^5x10 1.85 x 10^6	99.92
0.2M EMS -$Na_2S_2O_3$	-	1900	220	14	-	1.83 x 10^6x 10 1.83x10^7	99.2
0.4M EMS + $Na_2S_2O_3$	10	-	-	-	-	1 x 10^3 x 10 1 x 10^4	99.999
0.4M EMS - $Na_2S_2O_3$	90	13	-	-	-	1.1 x10^4x 10 1.1 x10^5	99.995

*Original titre of T4 phage before treatment was 2.27x10^9 PFU/ml.

- Sign indicating the absence of plaque.

The number 10 is dilution factor.

Table 2 : EMS induced mutagenesis in *frd*1 mutants of bacteriophage T4 under the extracellular treatment conditions.

Concentration of mutagenic treatment to *frd*1 mutant of T4	No. of Mutant plaques (halo⁻) on each plate				Titre (PFU/ml)	Mutational Frequency (M.F.)*
	10^{-1}	10^{-2}	10^{-3}	10^{-4}		
0.04M EMS + $Na_2S_2O_3$	-	-	5	-	$5 \times 10^{4} \times 10$ 5×10^{5}	2×10^{-4}
0.04M EMS - $Na_2S_2O_3$	-	-	6	-	$6 \times 10^{4} \times 10$ 6×10^{5}	2×10^{-4}
0.1 M EMS + $Na_2S_2O_3$	-	-	6	-	$6 \times 10^{4} \times 10$ 6×10^{5}	2×10^{-4}
0.1M EMS - $Na_2S_2O_3$	-	-	3	-	$3 \times 10^{4} \times 10$ 3×10^{5}	1×10^{-4}
0.2M EMS + $Na_2S_2O_3$	-	-	-	7	$7 \times 10^{5} \times 10$ 7×10^{6}	3×10^{-3}
0.2M EMS -$Na_2S_2O_3$	-	-	-	6	$6 \times 10^{5} \times 10$ 6×10^{6}	2.6×10^{-3}
0.4M EMS + $Na_2S_2O_3$	-	-	-	-	-	-
0.4M EMS - $Na_2S_2O_3$	-	-	-	-	-	-

*Original titre of T4 phage before treatment was 2.27×10^{9} PFU/ml.

- Sign indicating the absence of plaque.

The number 10 is dilution factor.

Table. 3 : Marker rescue mapping of isolated mutants obtained from dihydrofolate reductase defective T4 phage with pJSS10 clone carrying the whole *nrdB* gene.

*frd*1 *nrdB* double mutants	MR Mapping with pJSS10	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques*
AM5	315	6.3
AM11	2500	50
AM38	3100	62
AM49	2500	50
AM54	3000	60
AM57	4250	85
AM66	250	5
AM80	400	8
AM84	275	5.5
AM91	400	8
AM92	400	8
AM100	665	13.3
AM104	1800	36
AM106	750	15
AM111	2500	50
AM112	3500	70
AM119	500	10
AM122	550	11
AM124	1000	20
AM141	3500	70
AM149	750	15
AM152	4375	87.5
AM153	4500	90
AM154	500	10
AM155	4750	95
AM160	500	10
AM162	350	7
AM166	4500	90
AM168	800	16
AM171	225	4.5
AM174	240	4.8
AM186	290	5.8
AM190	345	6.9
AM192	900	18
AM284	600	12
AM288	300	6
AM293	475	9.5
AM294	500	10
AM297	3500	70
AM300	1250	25
AM306	4000	80
AM310	250	5
AM312	450	9

*frd*1 *nrdB* double mutants	M R Mapping with pJSS10	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques *
AM326	2500	50
AM334	500	10
AM342	1600	32
AM345	500	10
AM350	350	7
AM354	400	8
AM358	425	8.5
AM359	3000	60
AM378	150	3
AM382	300	6
AM387	450	9
AM396	600	12
AM397	250	5
AM399	4500	90
AM405	400	8
AM408	550	11
AM415	605	12.1
AM424	550	11
AM425	3250	65
AM429	2500	50
AM436	750	15
AM437	530	10.6
AM442	2800	56
AM443	325	6.5
AM445	1750	35
AM449	2000	40
AM454	350	7
AM473	590	11.8
AM481	550	11
AM482	780	15.6
AM484	2000	40
AM487	750	15
AM492	1000	20
AM496	280	5.6
AM505	425	8.5
AM506	935	18.7
AM510	2550	51
AM511	375	7.5
AM514	440	8.8
AM532	530	10.6
AM536	405	8.1
AM538	475	9.5
AM550	435	8.7
AM552	455	9.1
AM557	550	11
AM559	330	6.6

*frd*1 *nrdB* double mutants	M R Mapping with pJSS10	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques *
AM560	1250	25
AM561	550	10
AM563	1050	21
AM567	300	6
AM570	450	9
AM571	340	6.8
AM581	1875	37.5
AM583	550	11
AM594	1250	25
AM596	350	7
AM597	405	8.1
AM601	475	9.5
AM603	400	8
AM611	750	15
AM614	750	15
AM616	415	8.3
AM626	750	15
AM635	350	7
AM636	625	12.5
AM642	1250	25
AM643	1000	20
AM647	455	9.1
AM651	750	15
AM654	2250	45
AM662	500	10
AM667	215	4.3
AM668	2000	40
AM673	250	5
AM680	900	18
AM685	687	13.75
AM693	465	9.3
AM694	405	8.1
AM695	625	12.5
AM703	400	8
AM704	780	15.6
AM710	215	4.3
AM713	590	11.8
AM714	1300	26
AM715	800	16
AM717	2945	58.9
AM719	2250	45
AM720	175	3.5
AM721	330	6.6
AM722	250	5
AM723	1200	24
AM724	2375	47.5

*frd*1 *nrdB* double mutants	M R Maping with pJSS10	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques *
AM727	1000	20
AM730	2550	51
AM732	1500	30
AM733	200	4
AM735	155	3.1
AM744	275	5.5
AM750	1665	33.3
AM753	1250	25
AM756	290	5.8
AM757	250	5
AM759	650	13
AM760	1500	30
AM764	2685	53.7
AM765	750	15
AM767	500	10
AM769	580	11.6
AM773	290	5.8
AM774	310	6.2
AM775	2500	50
AM779	4150	83
AM781	1000	20

* Total number of plaques screened. 5000

Table 4: The *nrdB* candidate mutants isolated from EMS treated *frd*1 mutants phage that exhibited MR$^+$ with pJSS10 clone carrying whole *nrdB* gene

156 Candidate mutants that show MR$^+$ with pJSS10					
AM5	AM11	AM38	AM49	AM54	AM57
AM66	AM80	AM84	AM91	AM92	AM100
AM104	AM106	AM111	AM112	AM119	AM122
AM124	AM141	AM149	AM152	AM153	AM154
AM155	AM160	AM162	AM166	AM168	AM171
AM174	AM186	AM190	AM192	AM284	AM288
AM293	AM294	AM297	AM300	AM306	AM310
AM312	AM326	AM334	AM342	AM345	AM350
AM354	AM358	AM359	AM378	AM382	AM387
AM396	AM397	AM399	AM405	AM408	AM415
AM424	AM425	AM429	AM436	AM437	AM442
AM443	AM445	AM449	AM454	AM473	AM481
AM482	AM484	AM487	AM492	AM496	AM505
AM506	AM510	AM511	AM514	AM532	AM536
AM538	AM550	AM552	AM557	AM559	AM560
AM561	AM563	AM567	AM570	AM571	AM581
AM583	AM594	AM596	AM597	AM601	AM603
AM611	AM614	AM616	AM626	AM635	AM636
AM642	AM643	AM647	AM651	AM654	AM662
AM667	AM668	AM673	AM680	AM685	AM693
AM694	AM695	AM703	AM704	AM710	AM713
AM714	AM715	AM717	AM719	AM720	AM721
AM722	AM723	AM724	AM727	AM730	AM732
AM733	AM735	AM744	AM750	AM753	AM756
AM757	AM759	AM760	AM764	AM765	AM767
AM769	AM773	AM774	AM775	AM779	AM781

Table. 5 : Marker rescue mapping of isolated mutants obtained from dihydrofolate reductase defective T4 phage with pJBK1 clone carrying the intron of *nrdB* gene.

*frd*1 *nrdB* double mutants	M R Mapping with pJBK1	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques *
AM11	2500	50
AM38	3000	60
AM49	2450	49
AM54	200	4
AM57	235	5
AM111	500	10
AM112	300	6
AM141	500	10
AM152	3500	70
AM153	500	10
AM166	2000	40
AM326	2000	40
AM342	225	5
AM359	500	10
AM399	450	9
AM425	600	11
AM429	4200	84
AM442	625	12
AM445	500	10
AM484	600	11
AM487	500	10
AM492	600	12
AM668	300	6
AM717	250	5
AM719	550	11
AM724	300	6
AM730	500	10
AM753	2450	49
AM764	3150	63
AM775	2850	57
AM779	2750	55

* Total number of plaques screened. 5000

Table 6: The *nrdB* candidate mutants isolated from EMS treated *frd*1 mutant phage exhibited MR$^+$ with pJBK1 subclone carrying intron of *nrdB* gene.

31 Candidate mutants that show MR$^+$ with pJBK1			
AM11	AM38	AM 49	AM54
AM57	AM111	AM112	AM141
AM152	AM153	AM166	AM326
AM342	AM359	AM399	AM425
AM429	AM442	AM445	AM484
AM487	AM492	AM668	AM717
AM719	AM724	AM730	AM753
AM764	AM775	AM779	

Table 7: Marker rescue mapping of the isolated intron mutants with pJG106 clone carrying A+B region of the intron

*frd*1 *nrdB* double mutants	M R Mapping with pJG106	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques *
AM11	None	-
AM38	None	-
AM49	None	-
AM54	200	4
AM57	300	6
AM111	1250	25
AM112	None	-
AM141	1750	35
AM152	None	-
AM153	None	-
AM166	None	-
AM326	None	-
AM342	None	-
AM359	1500	30
AM399	1000	20
AM425	None	-
AM429	None	-
AM442	None	-
AM445	None	-
AM484	None	-
AM487	None	-
AM492	None	-
AM668	750	15
AM717	None	-
AM719	None	-
AM724	None	-
AM730	None	-
AM753	None	-
AM764	None	-
AM775	None	-
AM779	None	-

* Total number of plaques screened. 5000

Table 8 : The *nrdB* candidate mutants isolated from EMS treated *frd*1 mutant phage exhibited MR$^+$ with pJG106 clone carrying (A+B) region of intron.

7 Candidate mutants that show MR$^+$ with pJG106		
AM54	AM57	AM111
AM141	AM359	AM399
AM668		

Table 9: Marker rescue mapping of the isolated intron mutants with pJG108 clone carrying C+D region of the intron.

*frd*1 *nrdB* double mutants	M R Mapping with pJG108	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques *
AM11	2500	50
AM38	2500	50
AM49	1750	35
AM54	None	-
AM57	None	-
AM111	None	-
AM112	750	15
AM141	None	-
AM152	4500	90
AM153	750	15
AM166	1250	25
AM326	1250	25
AM342	500	10
AM359	None	-
AM399	None	-
AM425	1250	25
AM429	4000	80
AM442	1000	20
AM445	1000	20
AM484	1250	25
AM487	900	18
AM492	1250	25
AM668	None	-
AM717	450	9
AM719	1000	20
AM724	600	12
AM730	1000	20
AM753	2500	50
AM764	3750	75
AM775	2250	45
AM779	2250	45

* Total number of plaques screened. 5000

Table 10: The *nrdB* candidate mutants isolated from the EMS treated *frd*1 mutant phage exhibited positive MR with pJG108 clone carrying C+D region of intron.

24 Candidate mutants that show MR^+ with pJG108.			
AM11	AM38	AM49	AM112
AM152	AM153	AM166	AM326
AM342	AM425	AM429	AM442
AM445	AM484	AM487	AM492
AM717	AM719	AM724	AM730
AM753	AM764	AM775	AM779

Table 11: Marker rescue mapping of the isolated intron mutants with pJBS1 clone carrying region A of the intron.

*frd*1 *nrdB* double mutants	M R Mapping with pJBS1	
	No. of halo$^-$ plaques	% Frequency of halo$^+$ plaques *
AM54	250	5
AM57	315	8.3
AM111	600	12
AM141	750	15
AM359	300	6
AM399	500	10
AM668	450	9

* Total number of plaques screened. 5000

Table 12: The *nrdB* candidate mutants isolated from EMS treated *frd*1 mutant phage that exhibited positive marker rescue with pJBS1 clone carrying the region A of intron.

7 Candidate mutants that show MR$^+$ with pJBS1		
AM54	AM57	AM111
AM141	AM359	AM399
AM668		

* Total number of plaques screened. 5000

Table 13: Marker rescue mapping of the isolatd intron mutants with pJSR1 clone carrying region B of the intron.

*frd*1 *nrdB* double mutants	M R Mapping with pJSR1	
	No. of halo plaques	% Frequency of halo plaques *
AM54	None	-
AM57	None	-
AM111	None	-
AM141	None	-
AM359	None	-
AM399	None	-
AM668	None	-

* Total number of plaques screened. 5000

Table 14: Marker rescue mapping of the isolated intron mutants with pJSK7 clone carrying the region D of the intron.

*frd*1 *nrdB* double mutants	M R Mapping with pJSK7	
	No. of halo$^+$ plaques	% Frequency of halo$^+$ plaques *
AM11	None	-
AM38	None	-
AM49	None	-
AM112	900	18
AM152	None	-
AM153	1000	20
AM166	None	-
AM326	None	-
AM342	900	18
AM425	500	10
AM429	None	-
AM442	250	5
AM445	250	5
AM484	750	15
AM487	1250	25
AM492	550	11
AM717	250	5
AM719	1000	20
AM724	200	4
AM730	250	5
AM753	None	-
AM764	None	-
AM775	None	-
AM779	None	-

* Total number of plaques screened. 5000

Table 15: The candidate mutants isolated from the EMS treated *frd*1 mutant phage that exhibited positive marker rescue with the pJSK7 clone carrying region D of intron.

13 Candidate mutants that show MR+ with pJSK7		
AM153	AM342	AM442
AM445	AM484	AM487
AM492	AM112	AM425
AM717	AM719	AM724
AM730		

Table 16: The candidate mutants isolated form the EMS treated *frd*1 mutant phage that are mapped in the C region by default.

11 Candidate mutants that are mapped by default within C region		
AM11	AM38	AM49
AM152	AM166	AM326
AM429	AM753	AM764
AM775	AM779	

Table 17: Description of Oligonucleotides.

Probes	nt[a]	Specificity	Sequence (5'-3')[b]	Complement of [b,c]
ExII	22	Exon II	CACGTG ATG AACAGC TTC ACCT	1865-1886
ExI-ExII	25	Splice Junction	TATCTT TTG CGT. GTACCTTTAACTT	1189-1200.1799-1811
ExI-IVS	23	Exon I-Intron	TATCTTTTGCGT.AAAATGCGCCT	1189-1211
IVS-ExII	23	Intron-Exon II	TCGAA CATACG. GTA CCT TTAACT	1788-1810

[a]nt = Number of nucleotides in oligomer.
[b]Dots represent sequence junction.
[c]The numbering of nucleotides complementary to oligomer is given (numbering as Sjoberg B. -M. *et al*, 1986)

Table 18: Dot blot splicing assay of the mutants mapped within the region A of the intron of *nrdB* gene.

RNA preparation of mutants	ExII	ExI-ExII	ExI-IVS	IVS-ExII	Splicing defect site
wt	++	++	-	-	
AM54	++	-	+++	-	5'
AM57	++	-	+++	-	5'
AM668	++	++	+++	-	PD(5')
AM111	++	-	++	-	5'
AM141	++	-	+	-	5'
AM399	++	-	+	-	5'
AM359	++	-	+	-	5'

Hybridization signal : Undetectable. - : Weak. + : Moderate. ++ . Intense. +++
Partial Splicing defect. PD.

Table 19: Dot blot splicing assay of the mutants mapped within the region C of the intron of *nrdB* gene.

RNA preparation of mutants	ExII	ExI - ExII	ExI-IVS	IVS-ExII	Splicing defect site
wt	++	++	-	-	
AM11	++	-	+	++	5' & 3'
AM38	++	-	+	-	5'
AM49	++	-	+	-	5'
AM166	++	-	++	-	5'
AM326	++	-	++	-	5'
AM429	++	-	+++	-	5'
AM753	++	-	+++	+++	5' & 3'
AM764	++	-	+++	+++	5' & 3'
AM152	++	+++	++	-	PD(5')
AM775	++	-	+++	-	5'
AM779	++	++	++	+++	PD(5' & 3')

Hybridization signal : Undetectable. - ; Weak. + ; Moderate. ++ . Intense. +++
Partial Splicing defect. PD.

Table 20: Dot blot splicing assay of the mutants mapped within the region D of the intron of *nrdB* gene.

RNA preparation of mutants	ExII	ExI - ExII	ExI-IVS	IVS-ExII	Splicing defect site
wt	++	++	-	-	
AM112	++	-	+++	++	5' & 3'
AM153	++	-	+++	-	5'
AM342	++	-	-	++	3'
AM425	++	-	-	++	3'
AM442	++	+++	-	+++	PD(3')
AM445	++	-	-	++	3'
AM484	++	-	-	+++	3'
AM487	++	-	-	+++	3'
AM492	++	-	-	++	3'
AM717	++	+++	-	++	PD(3')
AM719	++	-	-	++	3'
AM724	++	-	-	++	3'
AM730	++	++	-	++	PD(3')

Hybridization signal : Undetectable. - : Weak. + : Moderate. ++ . Intense. +++
Partial Splicing defect. PD.

Table 21 : Mutants of region A showing recombination frequency with various clones of *nrdB* intron.

	% frequencies of halo$^+$ plaques with*		
*frd*1 *nrdB* double mutants	pJBK1	pJG106	pJBS1
AM54	4	4	5
AM57	5	6	8
AM668	6	15	9
AM399	9	20	10
AM359	10	30	6
AM111	10	25	12
AM141	10	35	15

*Total number of plaques screened, 5000

Table 22: Mutants of region C showing the recombination frequency with various clones of *nrdB* intron.

	% frequencies of halo$^+$ plaques with*	
*frd*1 *nrdB* double mutants	pJBK1	pJG108
AM166	40	25
AM326	40	25
AM49	49	35
AM753	49	50
AM11	50	50
AM779	55	45
AM775	57	45
AM38	60	50
AM764	63	75
AM152	70	90
AM429	84	80

*Total number of plaques screened, 5000

Table 23: Mutants of region D showing the recombination frequency with various clones of *nrdB* intron.

frd\ *nrdB* double mutants	% frequencies of halo$^+$ plaques with*		
	pJBK1	pJG108	pJSK7
AM342	5	10	18
AM717	5	9	5
AM112	6	15	18
AM724	6	12	4
AM153	10	15	20
AM730	10	20	5
AM445	10	20	5
AM487	10	18	25
AM719	11	20	20
AM425	11	25	10
AM492	12	25	11
AM484	11	25	15
AM442	12	20	5

*Total number of plaques screened, 5000

Figure 1 : EMS mutagenesis in *frd*1 mutants of bacteriophage T4.

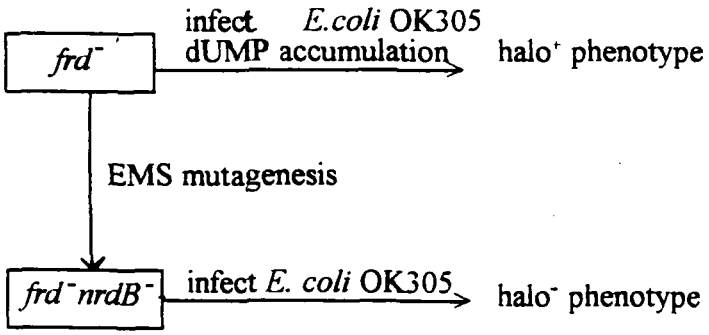

infect E.coli OK305
frd-
dUMP accumulation
halo+ phenotype
EMS mutagenesis
frd-nrdB-
infect E. coli OK305
halo- phenotype

Figure 2: Plaques showing halo$^+$ and halo$^-$ phenotypes.

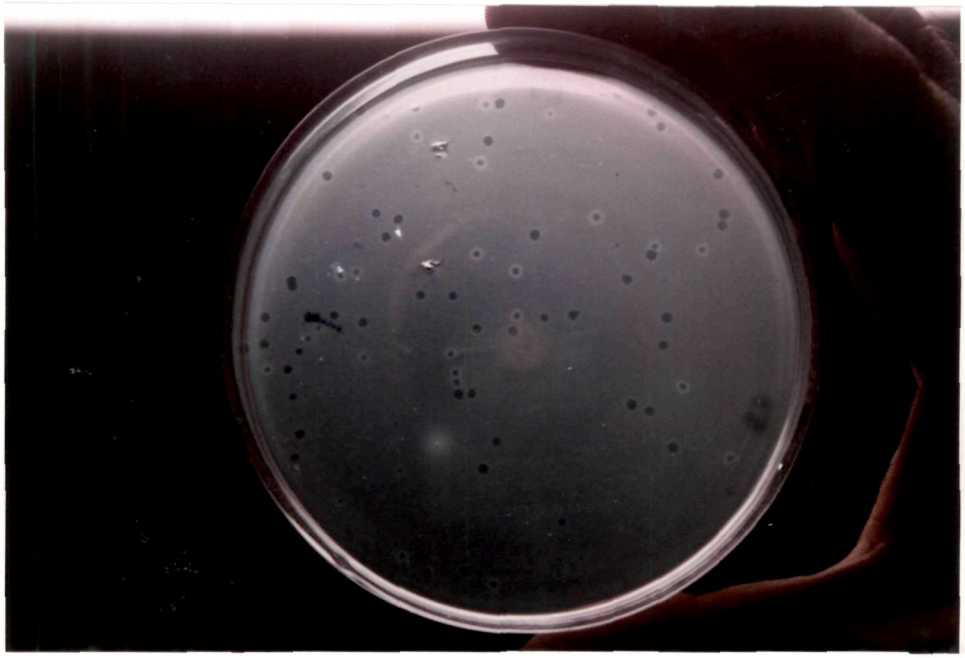


Figure 5: Strategy for the isolation and mapping of the *nrdB* intron mutants into various intron regions by marker rescue exploiting the white halo plaque phenotype.

Halo$^+$ character was assigned to these phage clones which formed at least 5% of the total plaques to be halo$^+$ otherwise the clone was considered as halo$^-$.

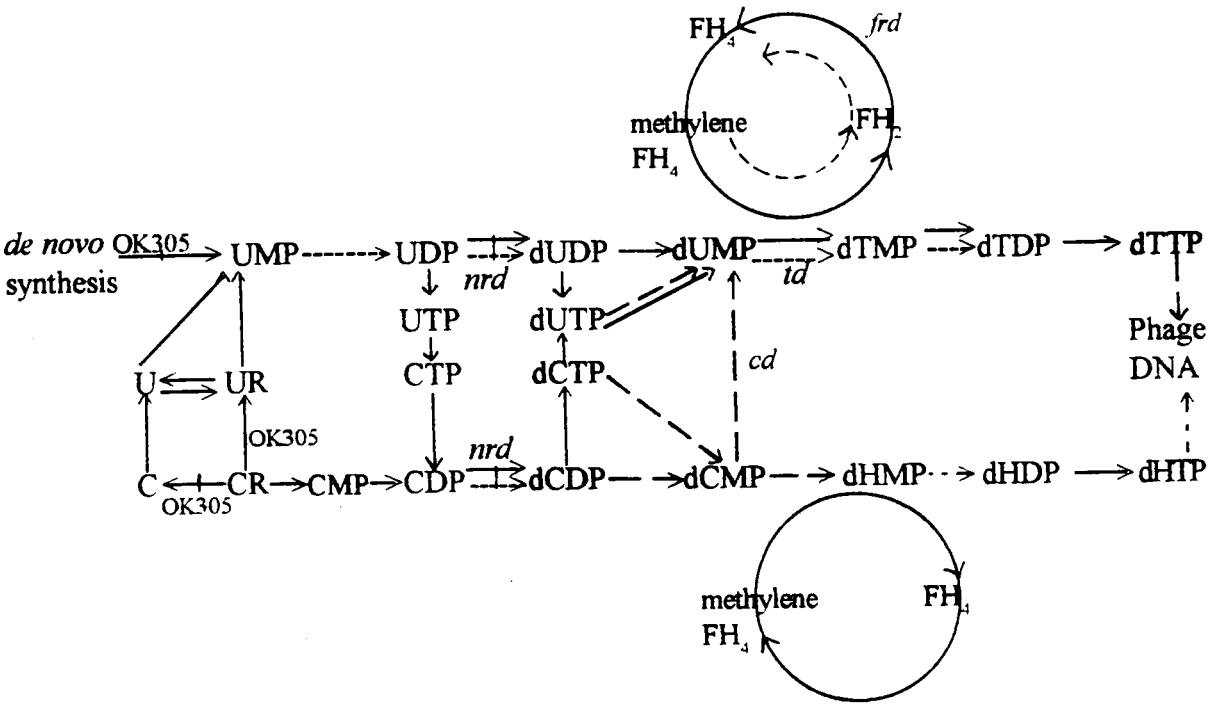

FH4
frd
methylene
FH4
FH2
de novo
synthesis
OK305
UMP
UDP
dUDP
dUMP
dTMP
dTDP
dTTP
nrd
td
UTP
dUTP
Phage
DNA
U
UR
CTP
dCTP
cd
OK305
nrd
C
CR
CMP
CDP
dCDP
dCMP
dHMP
dHDP
dHTP
OK305
methylene
FH4
FH4

Figure 4 : The pyrimidine nucleotide synthesis pathway. Solid lines represent bacterial (*E.coli*) enzymes. Dashed lines represent phage induced enzymes for which gene number or abbreviation is given, dHMP, dHDP, dHTP = hydroxymethyl cytidine nucleotide; FH_2 = dihydrofolate, FH_4 = tetrahydrofolate.

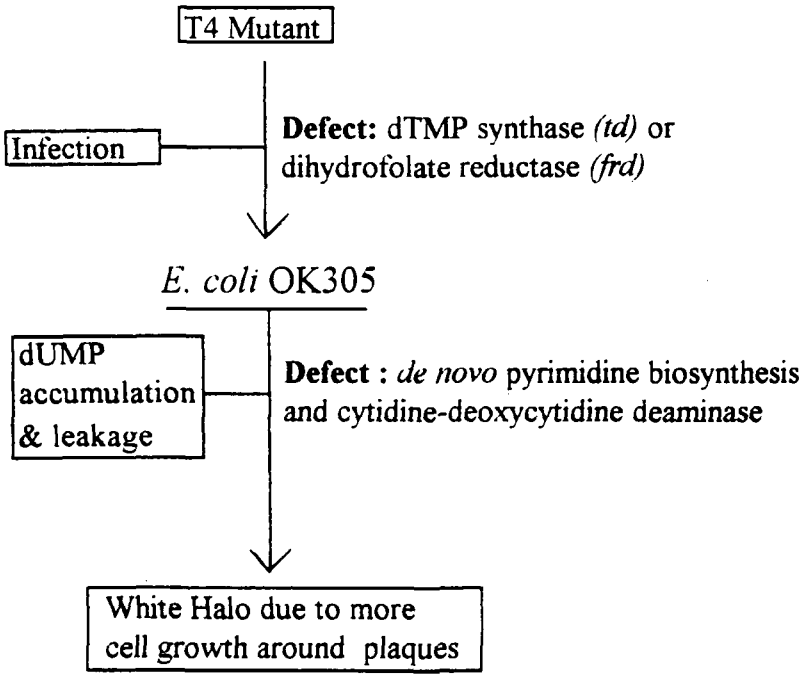

T4 Mutant
Infection
Defect: dTMP synthase (td) or dihydrofolate reductase (frd)
E. coli OK305
dUMP accumulation & leakage
Defect : de novo pyrimidine biosynthesis and cytidine-deoxycytidine deaminase
White Halo due to more cell growth around plaques

Figure 3 : The white halo plaque phenotype of bacteriophage T4.

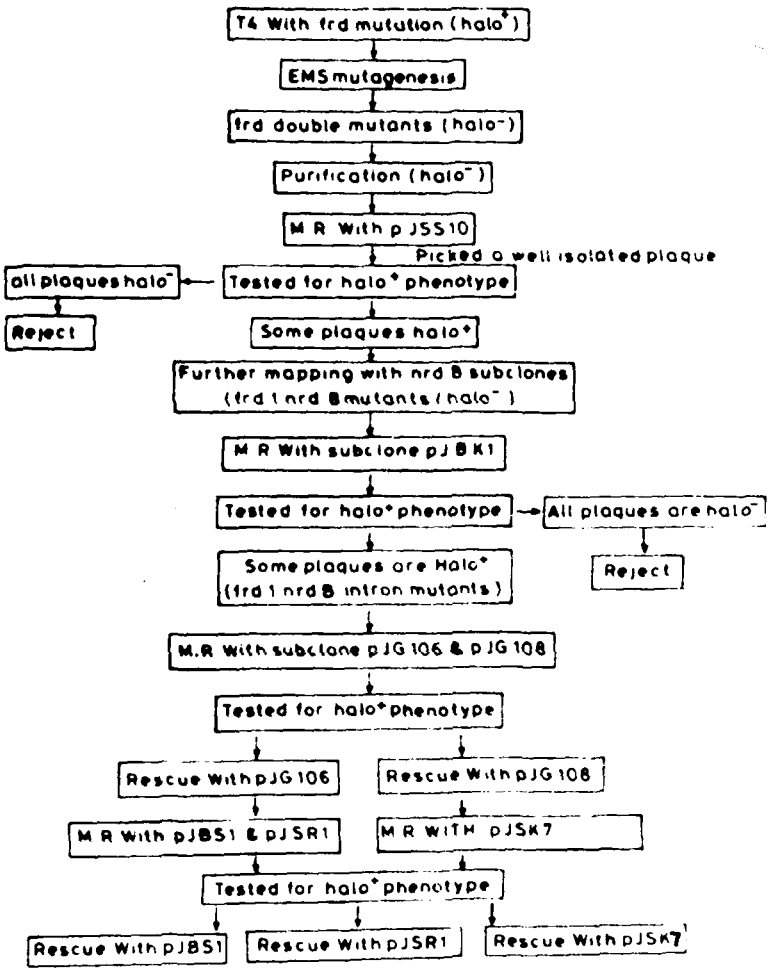

T4 With frd mutation (halo+)
EMS mutagenesis
frd double mutants (halo-)
Purification (halo-)
M R With p JSS10
Picked a well isolated plaque
all plaques halo-
Tested for halo+ phenotype
Reject
Some plaques halo+
Further mapping with nrd B subclones
(frd 1 nrd B mutants / halo-)
M R With subclone pJ B K1
Tested for halo+ phenotype
All plaques are halo-
Reject
Some plaques are Halo+
(frd 1 nrd B intron mutants)
M.R With subclone pJG 106 & pJG 108
Tested for halo+ phenotype
Rescue With pJG 106
Rescue With pJG 108
M R With pJBS1 & pJSR1
M R WITH pJSK7
Tested for halo+ phenotype
Rescue With pJBS1
Rescue With pJSR1
Rescue With pJSK7

Figure 6: A schematic diagram of the *nrdB* intron clones, pJBK1, complete intron carrying subclone (regions A + B + C + D) ; pJG106 (regions A + B), pJG108 (regions C + D), pJBS1 (region A), pJSR1 (region B), pJSK7 (region D) and region C divides the sequence covered by pJBK1 into smaller regions. *Bgl* II, *Sty* I, *Eco* R1, *Spc* I and *Kpn* I are the different restriction enzymes used for cloning and approximate locations are marked. Number in paranthesis indicates the number of mutation within that region.

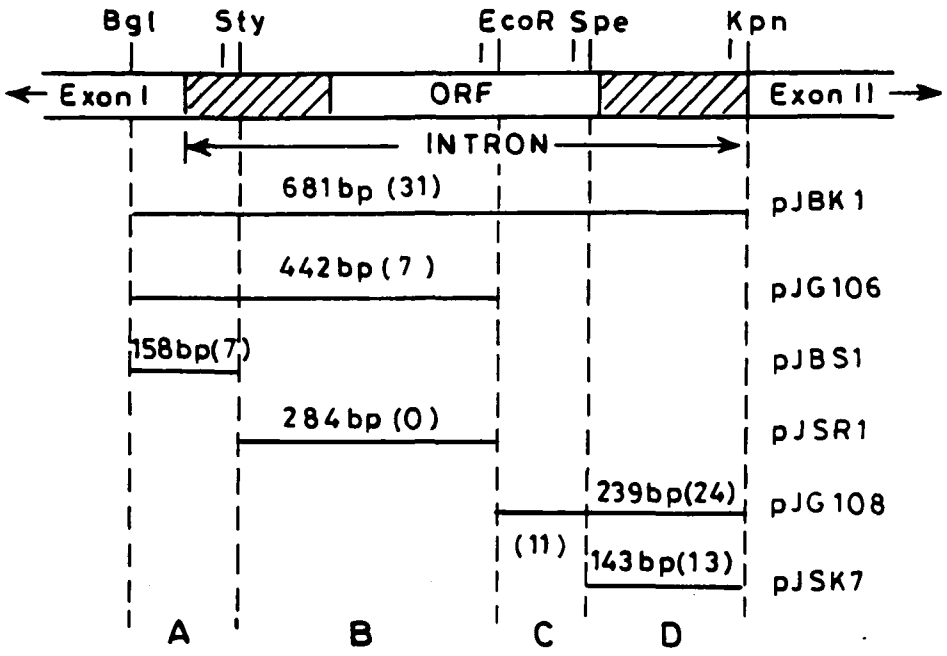

Bgl
Sty
EcoR
Spe
Kpn
Exon I
ORF
Exon II
INTRON
681bp (31)
pJBK1
442bp (7)
pJG106
158bp(7)
pJBS1
284bp (0)
pJSR1
239bp(24)
pJG108
(11)
143bp(13)
pJSK7
A
B
C
D

Figure 7: Dot blot splicing assay for the *nrdB* intron mutants of region A of the *nrdB* intron.

Probes, ExII, EXI-ExII, ExI-IVS, IVS-ExII correspond to Exon II, Splice junction, Exon I-Intron (5'cutting), and Intron-Exon II (3' cutting) oligonucleotides respectively.

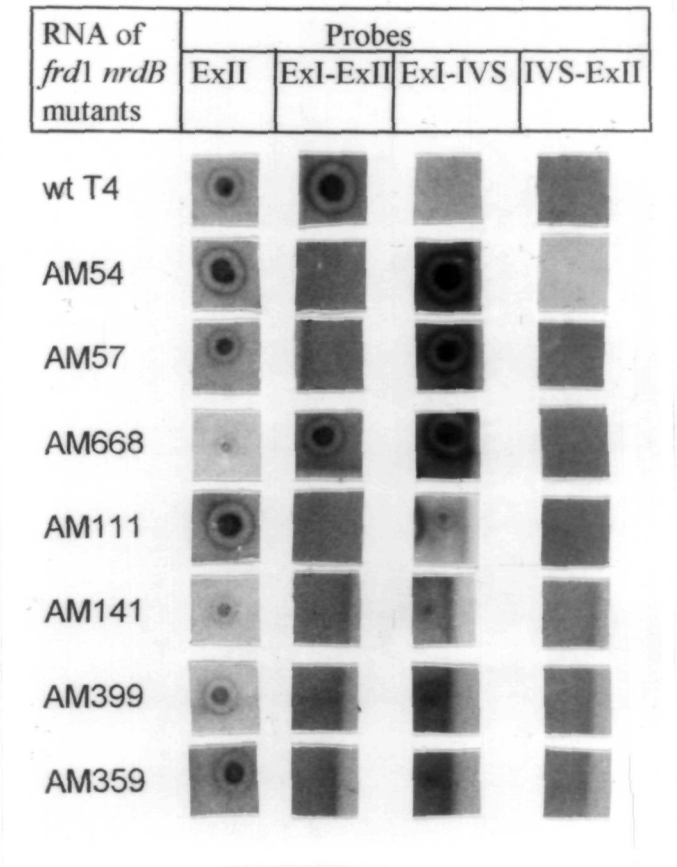

RNA of frd1 nrdB mutants
Probes
ExII
ExI-ExII
ExI-IVS
IVS-ExII
wt T4
AM54
AM57
AM668
AM111
AM141
AM399
AM359

Figure 8: Dot blot splicing assay for the *nrdB* intron mutants of region C of the *nrdB* intron.

Probes, ExII, EXI-ExII, ExI-IVS, IVS-ExII correspond to Exon II, Splice junction, Exon I-Intron (5'cutting), and Intron-Exon II (3' cutting) oligonucleotides respectively.

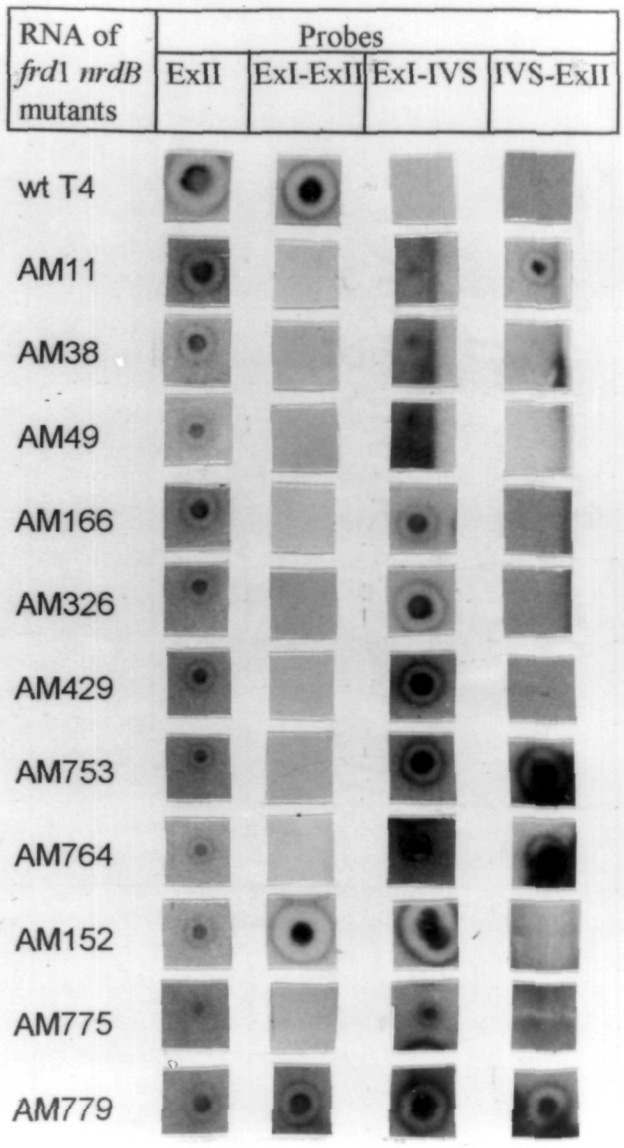

RNA of frd1 nrdB mutants
Probes
ExII
ExI-ExII
ExI-IVS
IVS-ExII
wt T4
AM11
AM38
AM49
AM166
AM326
AM429
AM753
AM764
AM152
AM775
AM779

Figure 9: Dot blot splicing assay for the *nrdB* intron mutants of region D of the *nrdB* intron.

Probes, ExII, EXI-ExII, ExI-IVS, IVS-ExII correspond to Exon II, Splice junction, Exon I-Intron (5'cutting), and Intron-Exon II (3' cutting) oligonucleotides respectively.

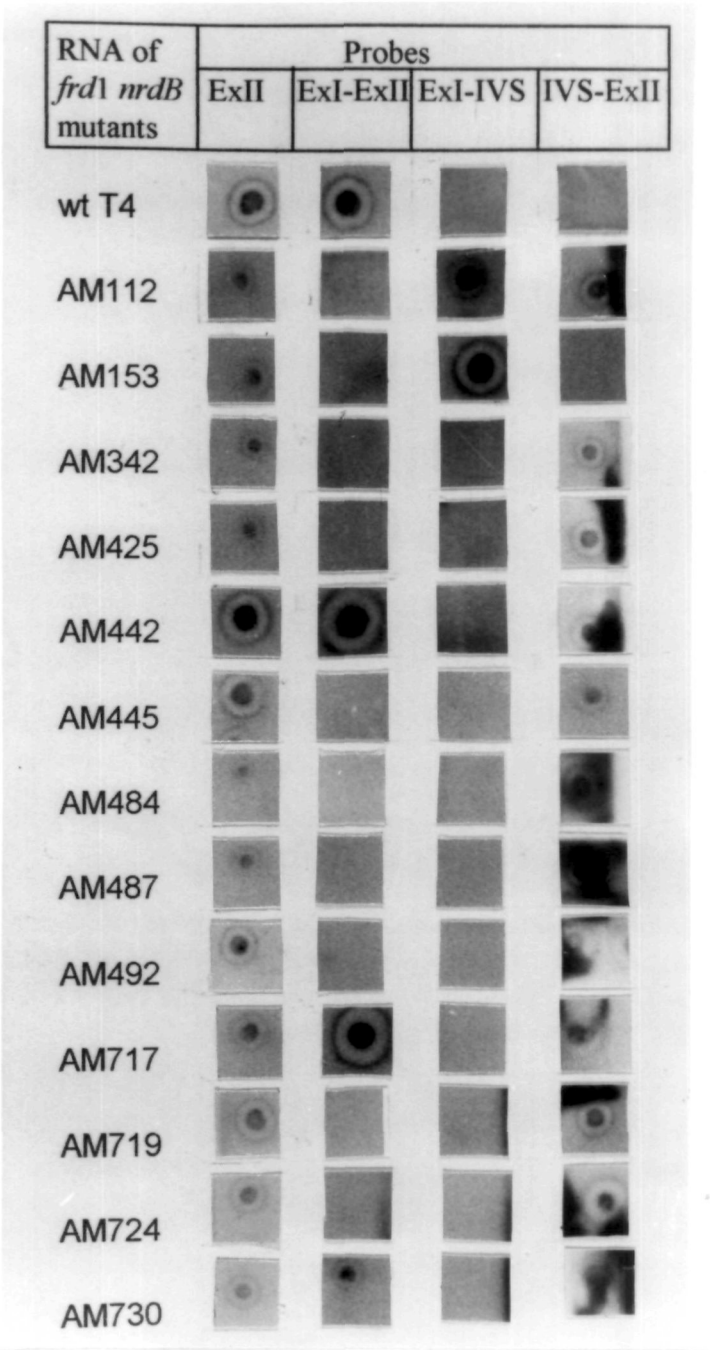
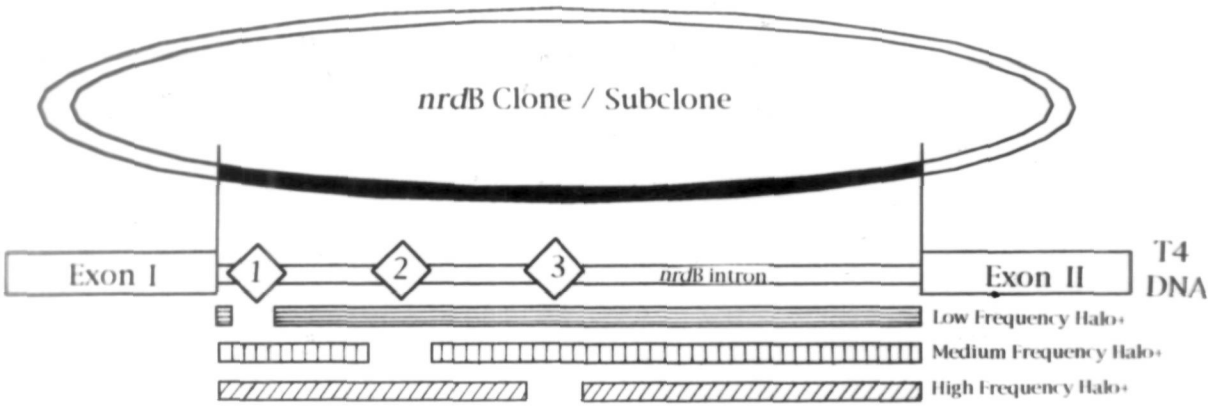

RNA of frd1 nrdB mutants
Probes
ExII
ExI-ExII
ExI-IVS
IVS-ExII
wt T4
AM112
AM153
AM342
AM425
AM442
AM445
AM484
AM487
AM492
AM717
AM719
AM724
AM730

Figure 10: A schematic view of the mechanism of marker rescue and its application in the prediction of genetic distances. The numbers 1, 2, and 3 within the rhombus show three different single point mutations within the *nrdB* intron for three independently isolated mutants. The plasmid contains the wt sequence for the *nrdB* clone being used for rescuing either one of the 1, 2, or 3 mutations (rhombus) within the *nrdB* intron of the T4 DNA. Depending upon the frequency of crossover, which is in direct relationship with the distance from the extremities of the wt insert, relative distance of that mutation can be estimated. Thus the mutation 3 will show a high degree of wt recombinants resulting in a high halo$^+$ frequency (diagonally shaded box), mutation 2 will show a medium degree (vertically shaded box) and mutation 1 will show frequency of halo$^+$ phenotype (horizontally shaded box).


nrdB Clone / Subclone
Exon I
1
2
3
nrdB intron
Exon II
T4
DNA
Low Frequency Halo+
Medium Frequency Halo+
High Frequency Halo+

CHAPTER IV: Isolation and mapping of intragenic and extragenic revertants of *frdA nrdB* double mutants of *nrdB* intron of bacteriophage T4

INTRODUCTION

It is known that the folded structure of group I intron is essential for its catalytic activity, as the three dimensional structure of a protein enzyme is important to its biological functions (Waring and Davies, 1984). Several approaches have been employed to reveal the important structural features for catalytic activity in group I introns. Comparative sequence analysis among group I introns of diverse organisms revealed highly conserved primary sequence elements, conserved secondary structure, and a predicted three-dimentional structure (Davies et al, 1982; Burke, 1988; Michel and Westhof, 1990). Site directed mutagenesis of the Tetrahymena rRNA gene and *in vitro* splicing reactions utilizing artificial substrates revealed the splicing mechanism and specific structural elements important for self splicing of the ribozyme of Tetrahymena (Cech, 1990; Pyle et al, 1990; Pyle and Cech, 1991). Basic features of the predicted three dimensional structure have been elucidated by mutants analysis *in vitro* and by the use of specifically positioned photochemical cross linking and affinity cleavage reagents (Burke et al, 1990; Wang and Cech, 1992; Wang et al, 1993).

All the T4 introns (*td, nrdB, nrdD*) are capable of self-splicing and contain the conserved structural features folded into a common secondary structure of group I introns (Waring and Davies, 1984; Shub et al, 1988; Belfort, 1990). The *td* intron is the best characterized and has contributed a great deal of information to the understanding of important structural features of the group I introns (Hall et al, 1987; Heuer et al, 1991). To examine the predicted secondary and tertiary structure of group I introns of T4 phage, random mutagenesis utilizing novel technique is very useful. The advantage of this approach is based on phenotypic screening methods (Lal and

Hall, 1993, 1997) and easy mapping of mutations due to the promiscuous recombination system of T4 (Broker and Doermann, 1975).

Single base changes in predicted RNA pairing regions which disrupt splicing, provide proof that a nucleotide is essential for the splicing reaction. Genetic verification for the involvement of a base pair can be demonstrated by the isolation of a pseudorevertants that harbours a compensatory mutation at the nucleotide predicted to base pair with the original mutant nucleotide. The double mutants would be predicted to regenerate a Watson-Crick base pair in an essential RNA helix, thus permitting intron RNA to fold and resulting in the restoration of some degree of autocatalytic splicing. Such experiments have been carried out using site specific mutagenesis to verify the existence of group I intron RNA pairing regions in Tetrahymena (Burke, 1988; Flor et al, 1989; Williamson et al, 1989; Couture et al, 1990) and yeast mitochondria (Weiss-Brummer et al, 1983; Holl et al, 1985).

In the previous chapter, we have described the isolation of 156 EMS induced *nrdB* mutants. Thirty one of them contained mutations which map in or near the *nrdB* intron. The isolation of second site revertants, which restore splicing activity of the *nrdB* intron mutants by introducing complementary mutations at second site, might significantly contribute to the understanding of secondary and tertiary interactions in the splicing process *in vivo*.

This chapter deals with the study of such revertants to gain an insight into the underlying process of autocleavage of pre-*nrdB* mRNA as well as the critical sites required for acquiring the biologically functional structure.

METHODS

EMS mutagenesis of *frd1nrdB* double mutants of phage T4 particles was performed (Fig.1) as described in chapter II.

After the EMS mutagenesis of *frd1nrdB* double mutants, we screened the phenotypic revertants for $nrdB^+$ marker, based on the halo phenotype. Here the revertants (i.e. $nrdB^+frd1^-$) would change their phenotype from $halo^-$ to $halo^+$ in the pyrimidine limited GPTG medium.

All the revertants were subjected to marker rescue mapping with the clones of *nrdB* gene and intron as well as subclones of various intronic regions, as described in the section of general materials and methods.

The RNA preparations of the revertants were made as described by Lal and Hall (1997) and were U.V. cross linked on to the Hybond nylon membrane. The DNA probes, as described in chapter III were used for dot blot hybridization assay (Belfort et al, 1985).

RESULTS

Isolation of second site revertants: Thirty one *frd1nrdB* double mutants as described in chapter III were subjected to EMS mutagenesis to obtain phenotypic revertants of *nrdB* intron mutants. The probability of getting true reversion is remote but the phenotypic reversion brought about by the suppressor mutation stands a significant chance. These second site or pseudorevertants could again be screened by means of the halo phenotype (Fig. 2). Reversion in the halo phenotype would occur presumably because of the acquisition of the normal biological activity of *nrdB* gene product.

This sort of normal biological activity would result because of the normal splicing of the *nrdB* intron and thus to give rise the normal coding template for the translation process. The role of typical secondary and tertiary structure for the self splicing of introns is well established. Moreover the process of conserved region in the group I introns for the splicing function has also been well documented. Our mutants mapped within the *nrdB* intronic region, would obviously be carrying mutation in the critical regions. The acquisition of the desired conformation would require either the change of the mutant base to the normal one or else the parent mutant site being unaltered the change might also take place at another site within the intron itself so that the conformation is compatible to splicing.

The mutants located within the region A of intron were subjected to undergo the second attempt of mutagenesis and was found that 6 *frd*1*nrdB* mutants showed their revertants. In the region C, 8 mutants displayed revertants and there were 9 mutants in the region D that exhibited the revertants (Table 1).

Mapping of revertants of *frd*1*nrdB* intron mutants with the pJSS10 clone: All the 23 revertants (Table 2) were subjected to the marker rescue mapping with the clone of *nrdB* gene (pJSS10) (Fig. 2). It was found that all the revertants of the *frd*1*nrdB* mutants showed positive marker rescue with this clone except AM49R which was not mapped within the *nrdB* gene (Table 3). This data indicates that AM49R was an extragenic revertant while the remaining revertants were intragenic (Table 4).

Mapping of revertants of *frd*1*nrdB* intron mutants with the pJBK1 clone: These 22 intragenic revertants were then subjected to marker rescue with pJBK1 clone carrying the intronic region of *nrdB* gene (Fig. 2). Four

revertants could exhibit the marker rescue with the clone indicating that the second mutational site must have been present in the intronic region (Table 5,6).

Finer mapping of revertants with the sub-clones of intron: The finer mapping of these four revertants (AM111R, AM429R, AM487R and AM724R) was performed with the various subclones of the intron.Only the AM111R displayed a positive marker rescue with the pJG106 clone carrying the region A+B of the intron. The remaining three revertants exhibited negative response, indicating that these revertants could not be mapped in A+B region (Table 7). Another set of marker rescue mapping was performed with the subclone pJBS1 for mapping the AM111R. It was confirmed, that this revertant must lie within the region A of intron (Table 8).

All the four revertants were subjected to marker rescue mapping with pJG108 also, and AM429R, AM487R, AM724R were found to be mapped in C+D region (Table 9).

An interesting result was observed with AM429R, whose primary mutation was mapped within the region C (Table 9) but on marker rescue mapping with pJSK7, the clone carrying D region of *nrdB* intron, showed positive response with this clone and hence the site of reversion must be present within the region D.

The other two revertants (AM724R, AM487R) which were carrying the primary mutation within D region could also mapped within the same subregion (Table 10).

Thus it is clear from our data that the four intronic mutations could be suppressed by the reversion within the intronic region itself which was quite clear from the suppression within the C region by the one having the second mutation in the D region.

Dot blot hybridization splicing assay: The hybridization of revertants was performed similarly as discussed in the chapter III. In this dot blot assay, synthetic oligonucleotide probes, designed in such a way that their binding to the complementary regions on the RNA is dependent on the presence of the spliced or unspliced *nrdB* RNA, were used as probes (as discribed in chapter III). The control probe ExII, binding to both premessage and ligated message, gave a signal of the same intensity with wild type. The probe ExI-ExII, was directed against the ligated splice junction. The probe ExI-IVS, was specific to the 5' pre-mRNA splice sites while probe IVS-ExII , specifically bind to the 3' pre-mRNA splice site. The EXII probe served as a positive control because it would bind to all spliced and unspliced *nrdB* RNA. None of the RNA preparations exhibited negative response with this probe (Fig. 3,4,5).

Dot blot splicing assay of the revertants of *frd1nrdB* mutants of the region A of *nrdB* intron: None of the revertants displayed positive binding with the probe ExI-ExII. It showed the splicing defective nature of all the revertants of this region. The probe ExI-IVS, bound to the revertants AM54R, AM57R, AM111R and AM141R whereas the revertants AM399R and AM359R did not show positive binding to this probe. The revertants AM399R and AM359R exhibited binding to the probe IVS-ExII and rest of the revertants did not show any binding to this probe (Fig. 3). This data revealed that revertants AM54R, AM57R, AM111R and AM141R have defect at their 5'cutting site whereas AM399R and AM359R have defect at their 3'cutting site

(Table 11).

Dot blot splicing assay of the revertants of *frd1nrdB* mutants of the region C of *nrdB* intron: None of the revertants except AM753R displayed any significant binding with the probe ExI-ExII. The probe ExI-IVS showed positive binding to AM49R whereas others displayed negative binding to this probe. The revertant AM49R did not show binding to the probe IVS-ExII whereas remaining revertants of this region exhibited positive binding to this probe (Fig. 4). This data implicated that revertant AM49R is defective at its 5'cutting site. But AM753R seems to be partial splicing defective at its 3' cutting site. The remaining revertants displayed the defects at 3'cutting site (Table 11).

Dot blot splicing assay of the revertants of *frd1nrdB* mutants of the region D of *nrdB* intron: Out of 9 revertants none of them exhibited positive binding with the probe ExI-ExII indicating the splicing defective point mutations (Fig. 5). The revertants, AM442R, AM445R, AM492R, and AM719R, showed positive binding to ExI-IVS and did not show binding to the revertants, AM112R, AM342R, AM487R, AM724R, and AM730R. The probe IVS-ExII bound to all of the revertants of this region except AM442R and AM719R which did not show any binding to this probe (Fig. 5). The revertants AM445R and AM492R showed the defect at both 5' and 3' cutting sites whereas AM442R and AM719R displayed the defect at 5'cuttig site. The remaining revertants exhibited the defects at their 3'cutting site (Table 11).

DISCUSSION

The ability of group I introns to catalyze their own splicing is mainly due to the precise folded structure. The folding of the intron RNA results in placing substrate site and catalytic site which are widely separated in the primary sequence in close proximity and activating phosphodiester bonds at splice sites (Waring and Davies, 1984). The catalytically important secondary and tertiary structures of group I introns have been studied with several approaches as described in the review of literature (Michel and Westhof, 1990). All the group I introns contain several conserved sequence elements, P, Q, R, S, which exhibit pair-wise complementarity. This sequence complementarity causes all group I introns to fold into strikingly similar structures involving specific pairing between the conserved sequences, P4, P6 and P7. This central core was proposed to bring the ends of the intron into close proximity, but deletion studies showed that some conserved paired regions can be deleted without affecting catalytic activity (Mattson et al, 1977). Some unpaired sequences between paired regions were also fairly conserved in primary sequences. It is thought that the paired regions do have a structural role to form secondary structure and the junction sequences have a functional role to catalyse self-splicing reactions. One way to understand the functional relevance of each sequence in paired and unpaired regions is to isolate splicing defective *nrdB* intron mutants (Lal and Hall, 1993) and their second site revertants which would have a compensatory mutation, thus restoring the splicing activity. Characterization of compensatory mutations involving sequencing and splicing assay was expected to give more comprehensive understanding of secondary and tertiary interactions for *nrdB* intron catalytic activity (Kwon et al, 1995; Lal and Hall, 1997).

We have, therefore, isolated 23 second site revertants of *frd*1*nrdB* intron mutations (Table 1). Some of the mutants, (AM11, AM153, AM425, AM484, AM666, AM717, AM764, and AM779) did not give any revertant. This could be explained on the basis of interactions involving other than regular Watson-Crick base pairs to form catalytically active structure, and therefore primary change would not be compensated by a second mutation in one of the interacting base (Kwon et al, 1995). Out of 23 revertants mapped within *nrdB* gene, only one revertant (AM45R) was found to be extragenic while the remaining 22 were all intragenic.

All the revertants were subjected to marker rescue mapping with pJBK1. This mapping resulted in the positive marker rescue with 4 out of 22 test revertants suggesting, that only a few compensatory mutations would lie within the intronic regions. Most of the *frd*1*nrdB* intragenic revertants were not mapped within the intron probably because these mutants might be present on either extreme ends of intron or a little bit toward the end of the exon I or start of exon II. Whereas the extragenic mutant that lies outside the *nrdB* gene, would probably disrupt the metabolic pathway of halo phenotype.

The extragenic mutation is very interesting since it could show the involvement of other gene product in folding the secondary structure to form a peculiar tertiary structure crucial for the splicing activity, or else the typical folding of the polycistronic RNA transcript containing the mutations in the *nrdB* and neighbouring genes.

Although several pre-mRNA containing group I introns are self splicing *in vitro*, there is compelling evidence that splicing *in vivo* requires proteins. In the case of yeast mitochondrial RNA splicing, these proteins have

been shown to include maturases encoded by ORF of its own intron (Garriga and Lambowitz, 1986).

Several trans-acting mutations have been identified, which block splicing of group I introns. Even though T4 introns are self splicing *in vitro*, the rate of splicing is significantly lower than that *in vivo*. In the case of *td* RNA, a non-physiologically high temperature (60°) is required to derive the self splicing reaction efficiently *in vitro*, compared to *in vivo* splicing at 37° or in the presence of an *E.coli* protein, synthesizing S-30 fraction, which suggests the participation of host and / or phage protein (s) (Belfort et al, 1986). Also the Neurospora CYT-18 protein, a tyrosyl tRNA synthetase, which functions in splicing group I introns in mitochondria, was reported to promote *in vitro* splicing of mutants of the distantly related bacteriophage T4 *td* intron (Mohr et al, 1992). According to this model, the enhancement of group I intron splicing by various protein factors, is brought about by their binding with the RNA to attain a typical conformation necessarily required for splicing. With this in mind, one of the goals of this study was to isolate extragenic revertants, restoring the splicing of defective *frd1nrdB* double mutants.

Dot blot splicing assay of our revertant AM49R suggested that the splicing proficiency of extragenic revertant was not restored. The unique characteristic of this revertant could be explained as follows. Since any mutation within the *frd/td* gene besides the defective *nrdB* gene would result in accumulation of dUMP, producing weak halos by compensating for the effect of leaky *nrdB* mutation. Thus AM49R would be an extragenic revertant, that could not affect the splicing of *nrdB* RNA, but affected the halo phenotype of the mutant, AM49. These results suggest that protein factors involved in the splicing of *nrdB* intron, may not be encoded by the

genome of T4 phage and hence these proteins must be encoded by its host chromosome.

Our dot blot results on the alteration in the splice site defect in the revertant from one cutting site to another are indicative of disruption of the secondary structure of intron (Table 11). In other words, the orientation of *defect had changed because of the disruption of the best conformation of* intron. For instance, AM399 has a defect at its 5'cutting site but after the second mutation in AM399, its revertant AM399R bears the defect at its 3' cutting site. Another mutant AM753 is splicing defective at both 5' and 3' (Table 19; chapter III) cutting site but its revertant AM753R has the defect at its 3'cutting site. This data suggest that the revertant, AM753R is restoring the splicing activity of primary mutant, AM753. Since primary mutant was completly splicing defective (Table 19; chapter III) whereas its revertant showed partial defect (Table 11; chapter IV) and hence we could suggest that the splicing proficiency of this extra-intronic, intragenic revertant has been restored to some extant.

Moreover, the mutant AM445 is defective at its 3'cutting site but its revertant AM445R has the defect at both the sites, i.e. 5' as well as 3' cutting sites.The mutant AM719 is also 3' splice site defective whereas its revertant AM719R has a defect at its 5'cutting site only (Table 11). These results, therefore, suggest that the cleavage of the intron involves precisely defined secondary structure acquiring a complicated tertiary structure which is vital for the splicing of group I type introns. Similar explaination was also given earlier (Lal and Hall, 1997).

Table 1 : Isolation of revertants from the *frd*1 *nrdB* double intron mutants of *nrdB* gene.

Primary Mutants	Region of primary mutations	Frequency of revertants	Revertants
AM54	A	4/5000	AM54R
AM57	A	6/5000	AM57R
AM668	A	0/5000	-
AM399	A	3/5000	AM399R
AM359	A	6/5000	AM359R
AM111	A	7/5000	AM111R
AM141	A	2/5000	AM141R
AM166	C	2/5000	AM166R
AM326	C	5/5000	AM326R
AM49	C	10/5000	AM49R
AM753	C	6/5000	AM753R
AM11	C	0/5000	-
AM779	C	0/5000	-
AM775	C	6/5000	AM775R
AM38	C	2/5000	AM38R
AM764	C	0/5000	-
AM152	C	2/5000	AM152R
AM429	C	3/5000	AM429R
AM342	D	7/5000	AM342R
AM717	D	0/5000	-
AM112	D	2/5000	AM112R
AM724	D	9/5000	AM724R
AM153	D	0/5000	-
AM730	D	7/5000	AM730R
AM445	D	5/5000	AM445R
AM487	D	10/5000	AM487R
AM719	D	6/5000	AM719R
AM425	D	0/5000	-
AM492	D	6/5000	AM492R
AM484	D	0/5000	-
AM442	D	1/5000	AM442R

(-) not detected.

Table 2: The *nrdB* candidate reverse intron mutants isolated from EMS treated *frd*1 *nrdB* double mutant of T4 phage.

23 Candidate revertant intron mutants isolated.			
AM54R	AM57R	AM399R	AM359R
AM111R	AM141R	AM166R	AM326R
AM49R	AM753R	AM775R	AM38R
AM152R	AM429R	AM342R	AM112R
AM724R	AM730R	AM445R	AM487R
AM719R	AM492R	AM442R	

Table 3 : Marker rescue mapping of the isolated revertants of *frd*1 *nrdB* double intron mutants with pJSS10 clone carrying *nrdB* gene.

		MR mapping with pJSS10
Revertants	Region of primary mutations	% frequency of halo⁻ plaques*
AM54R	A	56
AM57R	A	60
AM399R	A	80
AM359R	A	20
AM111R	A	40
AM141R	A	24
AM166R	C	40
AM326R	C	20
AM49R	C	-
AM753R	C	40
AM775R	C	36
AM38R	C	16
AM152R	C	40
AM429R	C	30
AM342R	D	10
AM112R	D	20
AM724R	D	50
AM730R	D	40
AM445R	D	20
AM487R	D	20
AM719R	D	30
AM492R	D	30
AM442R	D	60

*Total number of plaque screened, 5000.
(-) not detected

Table 4: The *nrdB* candidate reverse intron mutants isolated from the EMS treated *frd*1 *nrdB* double mutant exhibited MR$^+$ with pJSS10 clone carrying whole *nrdB* gene.

22 Candidate revertant intron mutants that show MR$^+$ with pJSS10*			
AM54R	AM57R	AM399R	AM359R
AM111R	AM141R	AM166R	AM326R
AM753R	AM775R	AM38R	AM152R
AM429R	AM342R	AM112R	AM724R
AM730R	AM445R	AM487R	AM719R
AM492R	AM442R		

*AM49R was not mapped with pJSS10. (Extragenic).

Table 5: Marker rescue mapping of the isolated revertants of *frd*1 *nrdB* double intron mutants with pJBK1 clone carrying *nrdB* intron.

		MR mapping with pJBK1
Revertants	Region of primary mutations	% frequency of halo⁻ plaques*
AM54R	A	-
AM54R	A	-
AM399R	A	-
AM359R	A	-
AM111R	A	12
AM141R	A	-
AM166R	C	-
AM326R	C	-
AM49R	C	-
AM753R	C	-
AM775R	C	-
AM38R	C	-
AM152R	C	-
AM429R	C	30
AM342R	D	-
AM112R	D	-
AM724R	D	7
AM730R	D	-
AM445R	D	-
AM487R	D	20
AM719R	D	-
AM492R	D	-
AM442R	D	-

*Total number of plaque screened, 5000.
(-) not detected

Table 6: The *nrdB* candidate reverse intron mutants isolated from EMS treated *frd*1 *nrdB* double mutant phage exhibited MR^+ with pJBK1 clone carrying *nrdB* intron.

4 Candidate revertant intron mutants that show MR^+ with pJBK1			
AM111R	AM429R	AM724R	AM487R

Table 7: **Marker rescue mapping of isolated reverse intron mutants which were mapped within the intron with pJG106 clone carrying A+B region of the intron.**

	MR mapping with pJG106		
Revertants	Region of primary mutations	% frequency of halo⁻ plaques*	Mapping position of revertants
AM111R	A	20	A+B
AM429R	C	-	C+D
AM724R	D	-	C+D
AM487R	D	-	C+D

*Total number of plaque screened, 5000.
(-) not detected

Table 8: Marker rescue mapping of isolated reverse intron mutants which were mapped within the A+B region of the intron with pJBS1 clone carrying A region of the intron.

	MR mapping with pJBS1		
Revertant	Region of primary mutation	% frequency of halo⁻ plaques*	Mapping position of revertant
AM111R	A	10	A

*Total number of plaques screened, 5000.

Table 9: Marker rescue mapping of the isolated reverse intron mutants which were mapped within the intron with pJG108 clone carrying C+D region of the intron.

	MR mapping with pJG108		
Revertants	Region of primary mutations	% frequency of halo⁻ plaques*	Mapping position of revertants
AM429R	C	20	C+D
AM724R	D	12	C+D
AM487R	D	18	C+D
AM111R	A	-	A

*Total number of plaque screened, 5000.
(-) not detected

Table 10: Marker rescue mapping of the isolated reverse intron mutants which were mapped within the C+D region of intron with pJSK7 clone carrying the D region of the intron.

	MR mapping with pJSK7		
Revertants	Region of primary mutations	% frequency of halo⁻ plaques*	Mapping position of revertants
AM429R	C	16	D
AM724R	D	10	D
AM487R	D	16	D

*Total number of plaques screened, 5000.

Table 11: Dot blot splicing assay of the *nrdB* revertants of *frd*1 *nrdB* mutants of *nrdB* intron.

RNA preparation of revertants	ExII	ExI-ExII	ExI-IVS	IVS-ExII	Splicing defect site
wt	+++	+++	-	-	
AM54R	+++	-	++	-	5'
AM57R	+++	-	+	-	5'
AM399R	+++	-	-	+++	3'
AM359R	+++	-	-	+++	3'
AM111R	+++	-	++	-	5'
AM141R	+++	-	++	-	5'
AM166R	+++	-	-	++	3'
AM326R	+++	-	-	++	3'
AM49R	+++	-	+	-	5'
AM753R	+++	++	-	+	PD(3')
AM775R	+++	-	-	+	3'
AM38R	+++	-	-	+++	3'
AM152R	+++	-	-	+++	3'
AM429R	+++	-	-	+++	3'
AM342R	+++	-	-	++	3'
AM112R	+++	-	-	+	3'
AM724R	+++	-	-	+	3'
AM730R	+++	-	-	++	3'
AM445R	+++	-	++	+	5'&3'
AM487R	+++	-	-	++	3'
AM719R	+++	-	+	-	5'
AM492R	+++	-	++	+++	5'&3'
AM442R	+++	-	++	-	5'

Hybridization signal : Undetectable, - ; Weak, + ; Moderate, ++ , Intense, +++
Partial Splicing defect, PD

Table 12 : The salient features of revertants.

Revertants	Region of primary mutation	Site of reversion*	Site of splice defect
AM54R	A	Extra intronic	5'
AM57R	A	,,	5'
AM399	A	,,	3'
AM359R	A	,,	3'
AM111R	A	A	5'
AM141R	A	Extra intronic	5'
AM166R	C	,,	3'
AM326R	C	,,	3'
AM49R	C	Extragenic	5'
AM753R	C	,,	PD(3')
AM775R	C	,,	3'
AM38R	C	,,	3'
AM152R	C	,,	3'
AM429R	C	D	3'
AM342R	D	Extra intronic	3'
AM112R	D	,,	3'
AM724R	D	D	3'
AM730R	D	Extra intronic	3'
AM445R	D	,,	5' & 3'
AM487R	D	D	3'
AM719R	D	Extra intronic	5'
AM492R	D	,,	5' & 3'
AM442R	D	,,	5'

*All extra-intronic revertants are intragenic

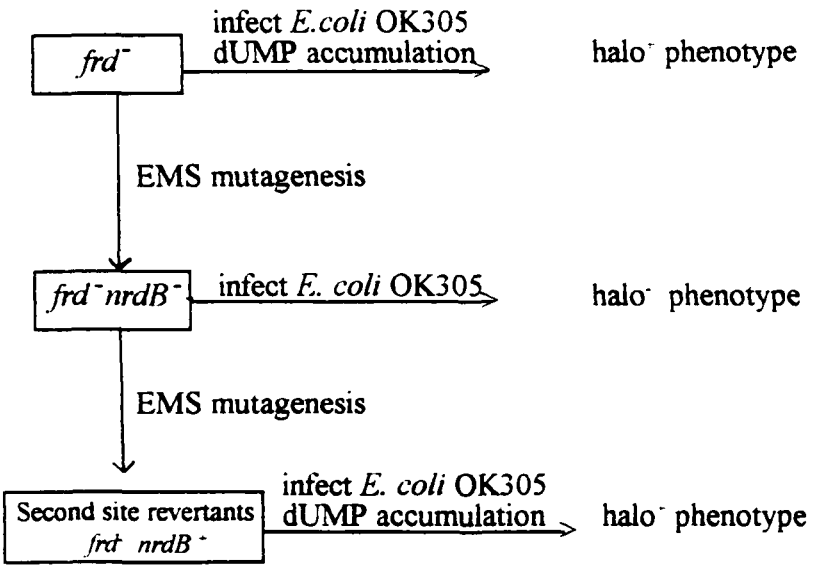

infect E.coli OK305
frd⁻
dUMP accumulation
halo⁻ phenotype
EMS mutagenesis
frd⁻ nrdB⁻
infect E. coli OK305
halo⁻ phenotype
EMS mutagenesis
infect E. coli OK305
Second site revertants
frd⁺ nrdB⁺
dUMP accumulation
halo⁻ phenotype

Figure 2: Strategy for the isolation and mapping of intragenic and extragenic revertants of *nrdB* intron mutants.

Halo$^+$ shows that the halo phenotype was clearly visible at minimum frequency of 5% of the plaques screened. Halo$^-$ indicates no plaques with halo phenotype (or less than 5%).

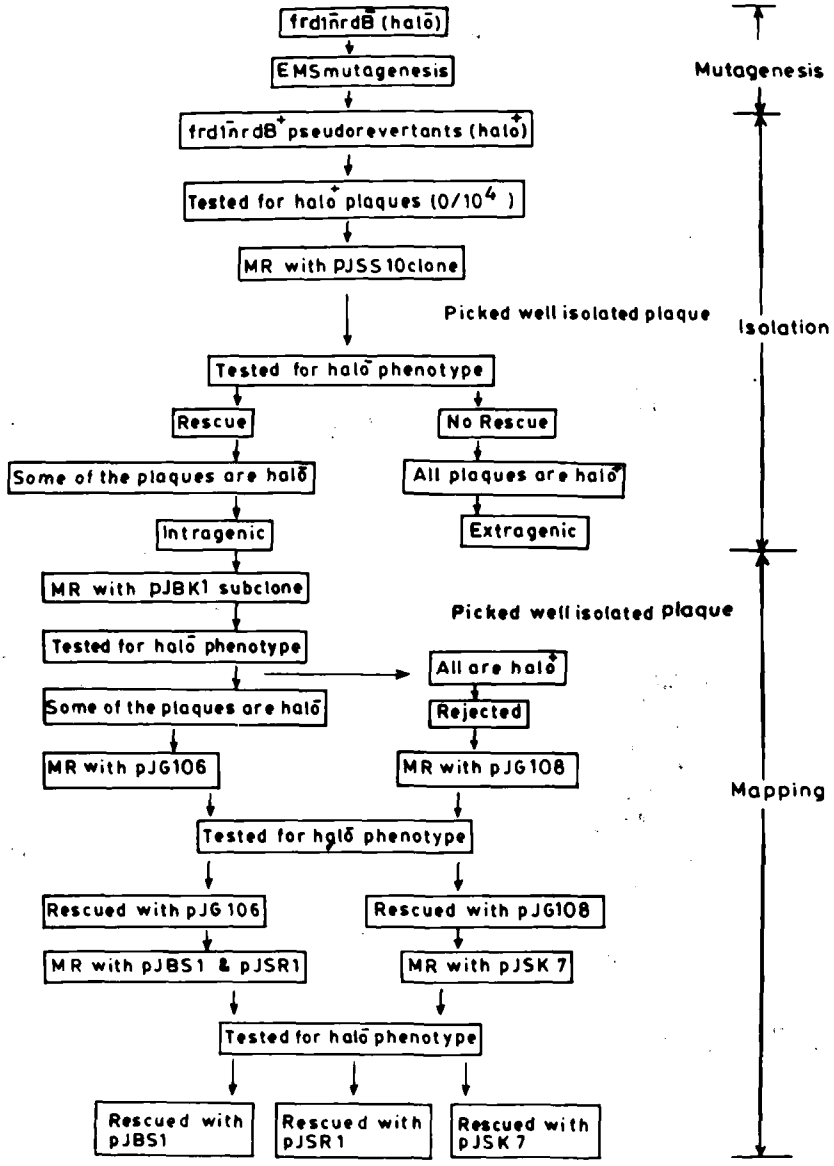

frd1⁻rdB⁻ (halō)
EMS mutagenesis
frd1⁻rdB⁺ pseudorevertants (halo⁺)
Tested for halo⁺ plaques (0/10⁴)
MR with pJSS10clone
Picked well isolated plaque
Tested for halō phenotype
Rescue
No Rescue
Some of the plaques are halō
All plaques are halo⁺
Intragenic
Extragenic
MR with pJBK1 subclone
Picked well isolated plaque
Tested for halō phenotype
All are halo⁺
Some of the plaques are halō
Rejected
MR with pJG106
MR with pJG108
Tested for hglō phenotype
Rescued with pJG106
Rescued with pJG108
MR with pJBS1 & pJSR1
MR with pJSK7
Tested for halō phenotype
Rescued with pJBS1
Rescued with pJSR1
Rescued with pJSK7
Mutagenesis
Isolation
Mapping

Figure 3: Dot blot splicing assay of the revertants of region A of the *nrdB* intron mutants.

Probes, ExII, EXI-ExII, ExI-IVS, IVS-ExII correspond to Exon II, Splice junction, Exon I-Intron (5'cutting), and Intron-Exon II (3' cutting) oligonucleotides respectively.

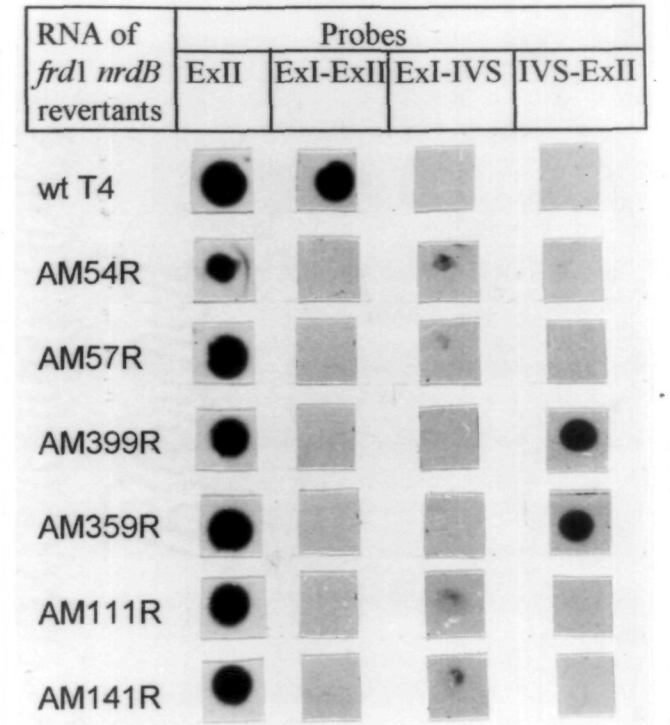

RNA of
frd1 nrdB
revertants
Probes
ExII
ExI-ExII
ExI-IVS
IVS-ExII
wt T4
AM54R
AM57R
AM399R
AM359R
AM111R
AM141R

Figure 4: Dot blot splicing assay of the revertants of region C of the *nrdB* intron mutants.

Probes, ExII, EXI-ExII, ExI-IVS, IVS-ExII correspond to Exon II, Splice junction, Exon I-Intron (5'cutting), and Intron-Exon II (3' cutting) oligonucleotides respectively.

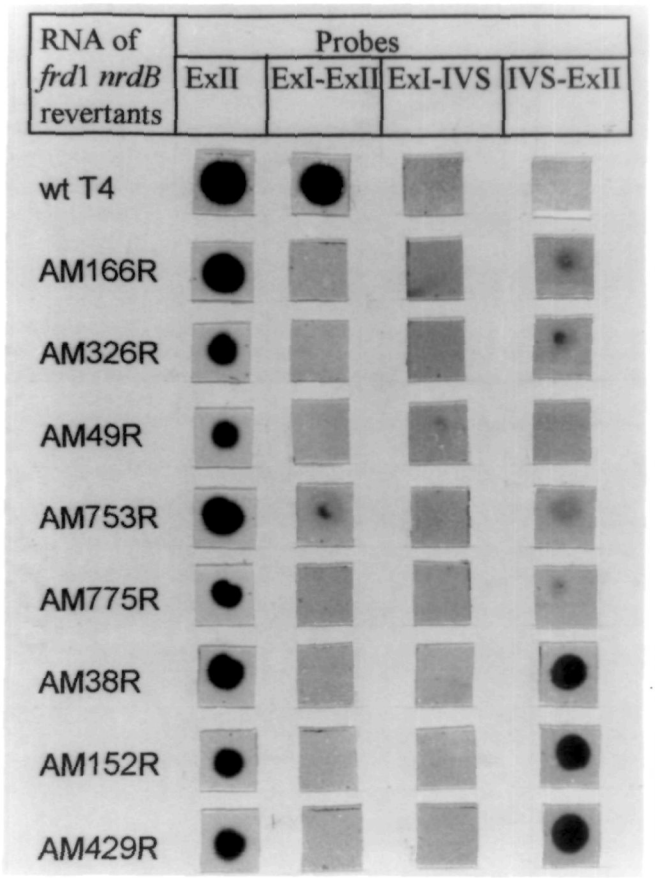
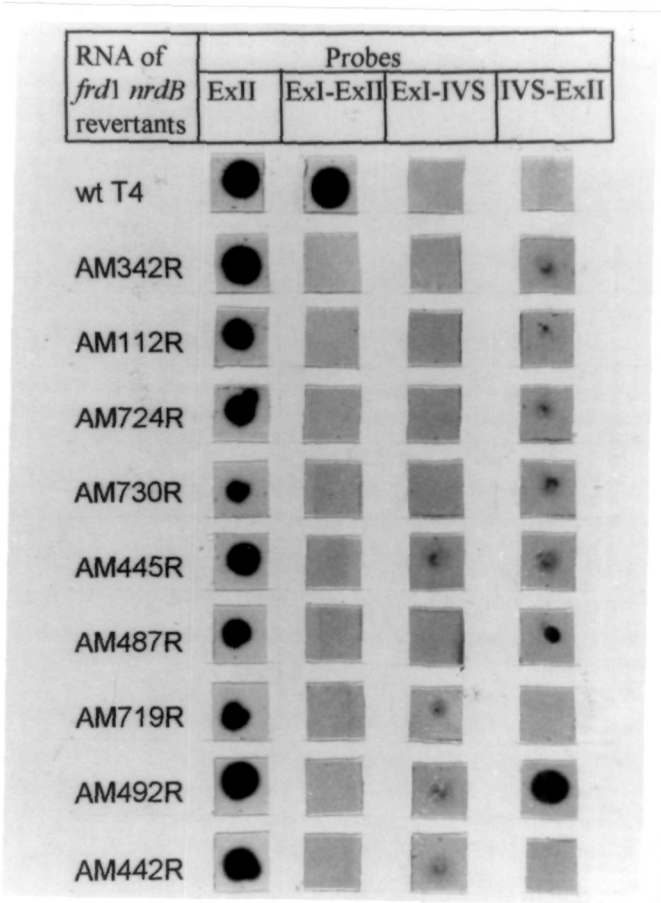

RNA of frd1 nrdB revertants
Probes
ExII
ExI-ExII
ExI-IVS
IVS-ExII
wt T4
AM166R
AM326R
AM49R
AM753R
AM775R
AM38R
AM152R
AM429R

Figure 5: Dot blot splicing assay of the revertants of region D of the *nrdB* intron mutants.

Probes, ExII, EXI-ExII, ExI-IVS, IVS-ExII correspond to Exon II, Splice junction, Exon I-Intron (5'cutting), and Intron-Exon II (3' cutting) oligonucleotides respectively.


RNA of frd1 nrdB revertants
Probes
ExII
ExI-ExII
ExI-IVS
IVS-ExII
wt T4
AM342R
AM112R
AM724R
AM730R
AM445R
AM487R
AM719R
AM492R
AM442R

CHAPTER V: General discussion

The biochemical and kinetic properties of group I intron auto-catalysis have been extensively investigated (Cech, 1990). Self-splicing of pre-mRNA *in vitro* has shown to occur in the absence of proteins or an external source of energy; only magnesium and guanosine are necessary. A common splicing mechanism has been demonstrated for several group I introns (Saldanha et al, 1993; Wittop-Koning and Schumperli, 1994). The energy requirement is obviated by trans-esterification mechanism in which the numbers of phosphodiester bonds formed and broken are the same. These introns thus contain all the structural features necessary for the RNA catalyzed reactions.

Ability of group I introns to perform autocatalytic splicing has now been exemplified by introns originating from phylogenetically diverse organisms. The discovery of introns in the T-even phages opened up bacterial genetic systems for the analysis of RNA splicing. Although quite different in primary sequence, a common feature of all group I introns is the presence, in a conserved order along the RNA, of several short consensus sequence elements that exhibit pairwise complementarity and guide folding of a core secondary structure that promote catalysis. Four of these elements are known as P,Q,R, and S. The internal guide sequence is also complementary to the 3' border of the 5' exon.

Additional sequence conservation of group I introns also includes a U at the 3' border of exon I and a G at the 3' border of the intron. Several of these conserved elements are known to be important because they are sites of cis-acting splicing mutations (Tabak and Grivell, 1986). The sequence complementarity between pair of elements may allow them to interact to form a complex higher order structure, consisting of helical segments, and

tertiary structures. This complex structure appears to be essential for splicing activity. These structures have been tested by compensatory mutations and by studies with enzymatic methods, chemical probes, NMR spectroscopy and computer algorithms (Wang and Cech, 1992; Allain and Varani, 1995 Gultyaev et al, 1995).

In the present study, taking advantage of phage genetics, we have tried to define the most crucial regions of the T4 *nrdB* intron that are critical to the group I splicing pathway. This was achieved by random mutagenesis , without any *a priori* knowledge or assumptions about regions functional in splicing. Thus by EMS - mutagenesis of the intact phage, and isolation and mapping of 156 *nrdB* mutations, main functional domains of the intron were delineated. None of the intron mutations mapped in its central 284 nucleotides stretch, indicating the dispensability of this region to splicing. On the other hand, clusters of mutations were mapped to the intron border regions, with at least 7 located within 158 nucleotides of the 5' end and at least 13 mapping within 143 nucleotides of the 3' end of the intron. The intron boundaries might be implicated as its functional splicing elements. The non-conserved structure is not irrelevent, it could provide a site to catalytic residues or to stabilize the overall structure of the intron. Moreover, it could also provide binding site for proteins that facilitate or regulate RNA splicing *in vivo* (Cech, 1988; Lal and Hall, 1997) as was demonstrated in certain yeast mitochondrial system (Lazowska et al, 1989; Mohr et al, 1992; Coetzee et al, 1994).

To understand the functional relevance of each sequence in paired and unpaired regions, the isolation of pseudo-revertants of *nrdB* intron mutants, having compensatory mutations vis-a-vis restoring the splicing activity would be of quite significance. We, therefore, isolated and

characterized *nrdB* intron mutants and their revertants employing *frd*1 mutants of T4 phage. Such mutations could also lead to more comprehensive understanding of secondary and tertiary interactions for the catalytic activity of *nrdB* intron.

Our study demonstrated that 3' region of ORF in the *nrdB* intron could also play a significant role in the splicing. In contrast, genetic studies on *td* and other group I introns demonstrated that intron ORF is not required for the active splicing conformation (Hall et al, 1987). This could probably be due to the small size of the ORF of *nrdB* intron (294 bases) compared to that of *td* intron (778 bases) which are looped out of the secondary structure models for RNA folding (Waring and Davies, 1984).

Comparing the dot-blot splicing data of the mutants with their second site revertants, we found that by changing the site of reversion, splicing defective site has also been changed (Table 18,19,20; Chapter III and Table 11; Chapter IV). For instance, mutant AM429 was primarily mapped in the region C while after reversion (AM429R) the mapping site has changed to lie in the region D (Table 12: Chapter IV). This site of reversion even within the short stretch of 239 nucleotides, has reversed the site of defect that is from 5' to 3' (Table 19; Chapter III).

These data suggest the probable involvement of some critical bases for integrity of the particular splice junction (5' or 3' cutting site).

Present study of the *nrdB* mutants of phage T4 and their revertants further leads us to conclude that the conserved sequences of *nrdB* intron that are important in pre-mRNA splicing, especially the extreme ends of *nrdB* intron, were more critical for autocatalysis. Moreover, the exon

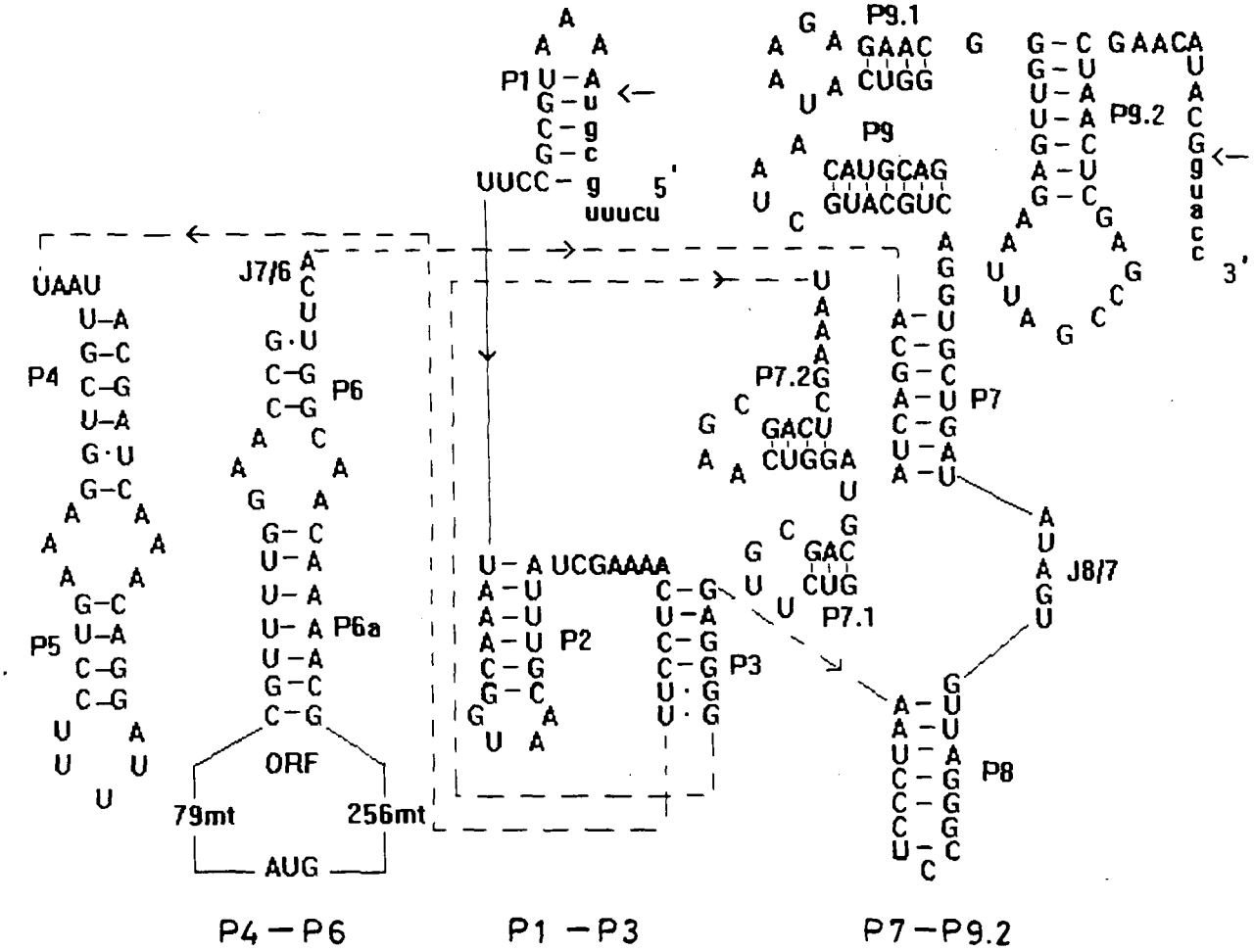
sequences immediately adjacent to the splice site were also important for splice site selection. Finally, the data also strongly suggest that the cleavage of intron might require a well defined secondary structure which on folding would probably acquire a typical tertiary structure involving the P9.0 and P10 loops. These secondary and / or tertiary structures are crucial for the splicing of *nrdB* intron like other group I type of introns.

In view of the above discussion, we would like to propose a plausible model to show the involvement of various domains of *nrdB* intron resulting in the formation of typical tertiary structure (Fig. 1).

The ORF region and neighbouring conserved region of P4-P6 domain stabilize the RNA structure essential for splicing. The catalytic core is formed by the folding of P1-P3 and P7-P9 domains. A conserved adenine in the catalytic core contributes to the stability of this arrangement and accepts a hydrogen bond from a specific 2'-OH of the P1 pairing region. Such backbone tertiary interactions might generally be important to the organization of RNA tertiary structure (Pyle et al, 1992).

This hypothetical model for the autocatalysis of *nrdB* intron is very close to that proposed earlier in case of Tetrahymena intron (Doherty and Doudna, 1997).

Figure 1: A proposed model for tertiary structure of *nrdB* intron involved in self splicing.
Solid lines show connectivities with each RNA domain while dashed lines indicate connectivities within the intact intron. Arrows superimposed on lines give 5' to 3' polarity. The 5' and 3' exon sequences are shown in lower case letters and both the sites are marked with arrows. Base-paired segments are designated as "P" and sections that join the two paired regions as "J".


P1
P2
P3
P4
P5
P6
P6a
P7
P7.1
P7.2
P8
P9
P9.1
P9.2
J7/6
J8/7
ORF
79mt
256mt
AUG
UAAU
UUCC
uuucu
UCGAAAA
CAUGCAG
GUACGUC
CUGGA
GACU
GAC
CUG
GAACA
GAAC
CUGG
P4 — P6
P1 — P3
P7 — P9.2

BIBLIOGRAPHY

Akins, R.A. and Lambowitz, A.M., (1987) Cell, 50 331-345.

Allain, F.H.-T. and Varani, G., (1995) J. Mol. Biol., 250 333-353.

Anziano, P.Q., Hanson, D.K., Mahler, H.R., and Perlman, P.S., (1982) Cell, 30 925-932.

Been, M.D. and Cech, T.R., (1985) Nucleic Acid Res., 13 8389-8408.

Been, M.D. and Cech, T.R., (1986) Cell,46 207-216.

Been, M.D., Barford, E.T., Burke, J.M., Price, J.V. and Tanner, N.K., (1988) Cold Spring Harbor Symp. Quant. Biol., 52 147-157.

Belfort, M., Pedersen-Lane, J., West, D. Ehrenman, K., Maley, G., Chu, F. and Maley, F., (1985) Cell,41 375-382.

Belfort, M, Pedersen-Lane, J., Ehrenman, K., Chu, F., Maley, G.F., Maley, F., Mc Pheeters, D.S. and Gold, L., (1986) Gene,41 93-102.

Belfort, M., Chandry, P.S. and Pedersen-Lane, P., (1987) Cold Spring Harb. Symp. Quant. Biol., 52 181-191.

Belfort, M., (1990) Annu.Rev.Genet., 24, 363-385.

Bell-Pedersen, D., Quirk, S.M., Aubrey, M. and Belfort, M., (1989) Gene 82 119-126.

Bell-Pedersen, D., Quirk, S.M., Clyman, J. and Belfort, M., (1990) Nucleic Acid Res., 18 3763-3770.

Berezney, R. and Coffey, D.S., (1977) J.Cell Biol., 73 616-620.

Birnboim, H.C. and Doly, J., (1980) Nucleic Acid Res., 7 1513.

Blum, B., Sturm, N.R., Simpson, A.M. and Simpson, L., (1991) Cell, 65 543-

550.

Brehm, S.L. and Cech, T.R., (1983) Biochemistry, 22 2390-2397.

Broker, T.R. and Doermann, A.H., (1975) Ann.Rev.Genet., 9 213-244.

Brookes, P. and Lawley, P.D., (1960) J.Chem.Soc., 339-545.

Brookes, P. and Lawley, P.D., (1962) J.Chem. Soc., 1348-1351.

Brown, M.D. Povinelli, C.M. and Hall, D.H., (1993) Biochem. Genet., 31 505-521.

Burke, J.M. and Rajbhandary, U.L., (1982) Cell, 31 509-520.

Burke, J.M., Irvine, K.D., Kaneko, K.J., Kerker, B.J., Oettgen, A.B., Tierney, W.M., Williamson, C.L., Zaug, A.J. and Cech,T.R., (1986) Cell, 45 167-176.

Burke, J.M., Belfort, M., Cech, T.R., Davies, R.W., Schweyen, R.J., Shub, D.A., Szostak, J.W. and Tabak, H.F., (1987) Nucleic Acid Research, 15 7217-7221.

Burke, J.M., (1988) Gene, 73 273-294.

Burke, J.M., Esherick, J.S., Burfeind, W.R. and King, J.L., (1990) Nature, 344 80-82.

Busch, H., Reddy, R., Rothblum, L. and Choi, Y.C., (1982) Ann.Rev.Biochem., 55 617-638.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R. and Daudona, J.A., (1996) Science, 273 1678-1685.

Cech, T.R., Zaug, A.J. and Grabowski, P.J., (1981) Cell, 27 487-496.

Cech, T.R., (1983) Cell, 34 713-716.

Cech, T.R., Tanner, N.K., Tinoco, I. Jr, Weir, B.R., Zuker, M. and Perlman,

P.S., (1983) Proc.Natl.Acad.Sci. (USA), 80 3903-3907.

Cech, T.R., (1986) Cell, 44 207-210.

Cech, T.R., (1887) Science, 236 1532-1539.

Cech, T.R., (1988) Gene, 73 259-271.

Cech, T.R., (1990) Annu. Rev. Biochem., 59 543-568.

Cech, T.R., Herschlag, D., Piccirilli, J.A. and Pyle, A.M., (1992) J.Biol. Chem., 267 17479-17482.

Chandry, P.S. and Belfort, M., (1987) Genes. Dev., 1 1028-1037.

Chandry, P.S., (1991) Ph.D. Thesis Albany Medical College Albany,NY.

Chiu, C.S., Tomich, P.K. and Greenberg, G.R., (1976) Proc Natl. Acad.Sci. (USA), 73 757-761.

Chiu, C.S., Cook, K.S. and Greenberg, G.R., (1982) J.Biol.Chem., 257 15087-15097.

Chowria, B.M., Berzal, H.A. and Burke, J.M., (1991) Nature, 354 320-322.

Christopher, D.A. and Hallick, R.B., (1989) Nucleic Acid Research, 17 7591-7607.

Chu, F.K., Maley, G.F., Maley, F. and Belfort, M., (1984) Proc. Natl. Acad. Sci.(USA), 81, 3149-3153.

Chu, F.K., Maley, G.F., Belfort, M. and Maley, F., (1985) J.Biol Chem., 260 10680-10688.

Chu, F.K., Maley, G.F., West, D.K., Belfort, M. and Maley F., (1986) Cell, 45 157-166.

Chu, F.K., Maley, G.F. and Maley, F., (1987) Biochemistry, 26 3050-3057.

Chu, F.K., Maley, G., Pedersen-Lane, J., Wang, A.M. and Maley, F., (1990) Proc.Natl. Acad.Sci. (USA), 87 3574-3578.

Coetzee, T., Herschlag, D. and Belfort, M., (1994) Genes and Dev., 8 1575-1588.

Couture, S., Ellington, A.D., Gerber, A.S., Cherry, J.M., Doudona, J. A. and Szostak, J.W., (1990) J.Mol.Biol., 215 345-358.

Daniel, C.J., Gupta, R. and Doolittle, W.F., (1985) J.Biol.Chem., 260 3132-3134.

Davies, R.W., Waring, R.B., Ray, J.A., Brown, T.A. and Scazzocchio,C., (1982) Nature, 300 719-724.

De La Salle, H., Jacq, C. and Slonimski, P.P., (1982) Cell, 28 721-732.

Doherty, E.A. and Daudona, J.A., (1997) Biochemistry, 36 3159-3169.

Doudona, J.A. and Szostak, J.W., (1989a) Nature, 339 519-522.

Daudona, J.A. and Szostak, J.W., (1989b) Mol.Cell.Biol., 9 5480-5483.

Doolittle, R.F., (1995) Proc.Natl.Acad.Sci.(USA), 92 2421-2423.

Drake, J.W. and Greening, E.O., (1970) Proc.Natl.Acad.Sci.(USA), 66 823-829.

Dropulic, B., Lin, N.H., Martin, M.A. and Jeang, K.T., (1992) J.Virology, 66 1432-1441.

Dujon, B., (1989) Gene, 82 91-114.

Ehrenman, K., Pedrsen-Lane , J., West, D., Herman, R., Maley, F. and Belfort, M., (1986) Proc. Natl. Acad. Sci.(USA), 83 5875-5879.

Ehrenman, K. Schroeder, R.M., Chandray, P.S., Hall, D.H. and Belfort,M.,

(1989) Nucleic Acid Res., 17 9147-9163.

Ferat, J.L. and Michel, F., (1993) Nature, 364 358-361.

Flor, P.J., Flanegan, J.B. and Cech,T.R., (1989) EMBO J., 8 3391-3399.

Forster, A.C., Jeffries, A.C., Sheldon, S.C. and Symon, R.H., (1987) Cold Spring Harb. Symp. Quant. Biol., 52 249-259.

Fox, G.E., Stackebrandt, E., Hespell, R.B., Gibson, J., Maniloff, J., Dyer, T.A. Wolfe, R.S., Balche, W.E., Tanner. R.S., Nagrum, L.J., Zalden, L.B.,Blakemore, R., Gupta, R., Lewis, B.J., Stahl, D.A., Luehrsen, K.R., Chen, K.N. and Woese, C.R., (1980), Science, 209 457-463.

Freese, E.B., (1961) Proc.Natl.Acad.Sci.(USA), 47 540-545.

Freese, E. and Freese, E.B., (1966) Radiat.Res.Suppl., 6 97-140.

Galloway-Salvo, J.L., Coetzee, T. and Belfort,M., (1990) J.Mol.Biol., 211 537-549.

Gampel, A. and Tzagoloff, A., (1987) Mol.Cell.Biol., 7 2545-2551.

Garriga, G. and Lambowitz, A.M., (1983) J.Biol.Chem., 258 14745-14748.

Garriga, G. and Lambowitz, A.M., (1984) Cell, 39 631-41 .

Garriga, G. and Lambowitz, A.M., (1986) Cell, 46 669-680.

Garriga, G., Lambowitz, A.M., Inoue, T. and Cech, T.R., (1986) Nature, 322 86-89.

Gott, J.M., Shub, D.A. and Belfort, M., (1986) Cell, 47 81-87.

Gott, J.M., Zeeh, A., Bell-Pedersen, D., Ehrenman, K., Belfort, M. and Shub, D.A., (1988) Gene Develop., 2 1791-1799.

Goodrich-Blair, H.A., Gott, J.M., Xu, M.Q., Scarlato, V. and Shub, D.A.,

(1989) Symp.Mol.Cell.Biol., 94 59-66.

Goodrich-Blair, H.A., Scarlato, V., Gott, J.M, Xu,M.Q. and Shub, D.A., (1990) Cell, 63 417-424.

Green, D.M. and Krieg, D.R., (1961) Proc.Natl.Acad.Sci. (USA), 47 64-72.

Green, M.R., (1986) Annu.Rev Genet., 20 671-708.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S., (1983) Cell, 35 849-857.

Gultyaev, A.P., Van Batenburg, F.H.D. and Pleij, C.W.A., (1995) J. Mol. Biol., 250 37-51.

Guo, Q. and Lambowitz, A.M., (1992) Genes & Dev., 6 1357-1372.

Hall, D.H., Povinelli, C.M., Ehrenman, K., Pedrsen-Lane , J., Chu, F. and Belfort, M., (1987) Cell, 48 63-71.

Hall, D.H., Tessman, I. and Karlstrom, O., (1967) Virology, 31 442-448.

Hampel, A. and Tritz, R., (1989) Biochemistry, 28 4929-4933.

Hampel, A., Tritz, R., Hicks, M. and Cruz, P., (1990) Nucleic Acid Res.18 299-304.

Herschlag, D. and Cech T.R., (1990a) Biochemistry, 29 10159-10171.

Herschlag, D. and Cech, T.R., (1990b) Nature, 344 405-409.

Heuer, T.S., Chandry, P.S., Belfort, M., Calender, D.W. and Cech, T.R., (1991) Proc.Natl. Acad. Sci.(USA), 88 1105-1109.

Hicke, B.J., Christian, E.L. and Yarus, M., (1989) EMBO J., 8 3843-3851.

Hickey, D.A., Benkel, B.F. and Abukashawa, S.M., (1989) J.Theor.Biol.,

137 41-53.

Hilfinger, J.M. and Ping, H., (1994) Gene, 142 55-60.

Holl, J., Rodel, G. and Schweyen, R.J., (1985) EMBO J., 4 2081-85.

Inoue, T., Sullivan, F.X. and Cech, T.R., (1985) Cell, 43 431-440.

Inoue, T., Sullivan, F.X. and Cech, T.R., (1986) J.Mol.Biol., 189 143-65.

Jacquier, A., (1990) Trends Biochem.Sci., 15 351-354.

Junn, E. and Kang, C., (1996) Ganet.Anal., 13 1-7.

Kaine, B.P., Gupta, R. and Woese, C.R., (1983) Proc.,Natl.Acad.Sci. (USA), 80 3309-3312.

Kim, S.H. and Cech, T.R., (1987) Proc.Natl.Acad.Sci.(USA), 84 8788-8792.

Kister, K.P. and Eckert, W.A., (1987) Nucleic Acid Res., 15 1905-1921.

Kjems, J. and Garrett, R.A., (1985) Nature, 318 675-677.

Koguma, T., Ohkawa, J., Mori, K., Fujita, S., Nishikawa, S., Iwakura, M. and Taire, K., (1995) Nucleic Acid .Symp. Ser, 34 221-222.

Koizumi, M., Ozawa, Y., Yagi, R., Wishigaki, T., Kaneko, M., Oka, S., Kimura, S., Iwamoto, A., Komatsu, Y. and Ohtsuka, E., (1995) Nucleic Acid Symp. Ser., 34 125-126.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech,T.R. (1982) Cell, 31 147-157.

Kwon, H.Y., Lal, S.K. and Hall, D.H., (1995) Nucleosides and Nucleotides, 14 1811-1821.

Lal, S.K. and Hall, D.H., (1993) Biochem. Biophys.Research Comm., 196 943-949.

Lal, S.K. and Hall, D.H., (1997) Biochem. Biophys. Acta, 1350 89-97.

Lambowitz, A.M., (1989) Cell, 56 323-326.

Lambowitz, A.M. and Perlman, P.S., (1990) Trends Biochem. Sci., 15 440-444.

Lawley, P.D. and Brookes, P., (1961) Nature, 192 1081-1082.

Lazowska, J., Claise, M., Gargouri, A., Kotylak, Z., Spyridakis, A. and Slonimski, P.P., (1989) J.Mol.Biol., 205 275-289.

Lehman, N. and Joyce, G., (1993) Nature, 361 182-185.

Lerner, M.R. and Steitz, J.A., (1981) Cell, 25 298-307.

Lieber, A., He, C.Y., Polyak, S.T., Gretch, D.R., Barr, D. and Kay, M.A., (1996) J. Virology., 70 8782-8791.

Loveless, A., (1958) Nature, 181 1212-1213.

Maniatis, T. and Reed,R., (1987) Nature, 325 673-678.

Mattson, T., Van Houwe, G., Bolle, A., Selzer, G. and Epstein, R., (1977) Mol.Gen.Genet., 154 319-326.

Menger, M., Tusch , T., Eckstein, F. and Porschke, D., (1996) Biochemistry, 35 14710-14716.

Michel, F., Jacquier, A. and Dujon, B., (1982) Biochimie., 64 867-881.

Michel, F. and Dujon, B., (1983) EMBO J., 2 33-38

Michel, F. and Cumming, D.J., (1985) Curr.Genet., 10 69-79.

Michel, F., Hanna, M., Green, R., Bartel, D.P. and Szostak,J.W., (1989a) Nature, 342 391-395.

Michel, M., Umesono, K. and Ozeki, H., (1989b) Gene, 82 5-30.

Michel, F. and Westhof,E., (1990) J.Mol.Biol., 216 585-610.

Mitchell, A., Richard, T. and Hampel, A., (1992) Gene, 122 85-90.

Moen, L.K., Howell, M.L., Lasser, G.W, and Mathews,C.K., (1988) J.Mol.Recognit., 1 48-57.

Mohr, G., Zhaug, A.X., Gianelos, J.A., Belfort, M. and Lambowitz, A.M., (1992) Cell, 69 483-494.

Morrisay, J.P. and Tollervey, D., (1995) Trends Biochemical Science, 20 78-82.

Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A., (1986) Annu.Rev. Biochem., 55 1119-1150.

Pal, B.C, (1962) Biochemistry, 1 558-562.

Partono, S. and Lewin, A.S., (1988) Mol.Cell.Biol., 8 2562-2571.

Pedersen-Lane, J. and Belfort, M., (1987) Science, 237 182-184.

Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A. and Cheng,H.L., (1986) Cell, 44 213-223.

Perlman, P.S. and Butow, R.A., (1989) Science, 237 182-184.

Perrault, J.P., Wu, T., Cousineau, B., Ogilvie, K.K. and Cedergren,R., (1990) Nature, 344 565-567.

Piccirilli, J.A., Vyle, J.S., Caruthers, M.H. and Cech, T.R., (1993) Nature, 361 85-88.

Poter, S., Durovic, P. and Dennis, P.P., (1995) Science, 268 1056-1060.

Price, J.V., Engberg, J. and Cech, T.R., (1987a) J.Mol. Biol., 196 49-60.

Price, J.V., Sullivan, F.X. and Cech, T.R., (1987a) J.Mol.Biol., 196 217-221.

Price, J.V. and Cech,T.R., (1988) Genes Dev., 2 1439-1447.

Pyle, A.M., Mc Swiggen, J.A. and Cech, T.R.,(1990) Proc. Natl. Acad. Sci. (USA), 87 8187-8191.

Pyle, A.M. and Cech, T.R., (1991) Nature, 350 628-631.

Pyle, A.M., Murphy, F.L. and Cech, T.R., (1992) Nature, 358 123-128.

Quirk, S.M., Bell-Pedersen, D. and Belfort, M., (1989a) Cell, 56 455-465.

Quirk, S.M., Bell-Pedersen, D., Tomaschewski, J., Ruger, W. and Belfort, M., (1989b) Nucleic Acid Res., 17 301-314.

Reinhold-Hurek, B. and Shub, D.A., (1992) Nature, 357 173-176.

Rochaix, J.D., Rahire, M. and Michel, F., (1985) Nucleic Acid Res., 13 975-984.

Rossi, J.J., (1995) TIBTECH., 13 301-306.

Saldanha, R., Mohr, G., Belfort, M. and Lambowitz, A.M., (1993) FASEB J., 7 15-24.

Sarver, N., Cantin, E.M., Chang, P.S., Zaia, J.A., Ladne, P.A., Stephens, D.A. and Rossi, J.J., (1990) Science, 247 1222-1225.

Scaringe, S.A., Francklyn, C. and Usman, N., (1990) Nucleic Acid Res., 18 5433-5441.

Scazzocchio, C., (1989) Trends Genet., 5 168-172.

Schmelzer, C. and Schweven, R.J., (1986) Cell, 46 557-565.

Shapiro, R. and Klein, R.S., (1966) Biochemistry, 5 2358-62.

Sharp, P.A., (1985) Cell, 42 397-405.

Sharp, P.A., (1987) Science, 235 766-771.

Shub, D.A., Gott, J.M., Xu, M.-Q., Franz Lang, B., Michel, F., Tomaschenski, J., Pedersen-Lane, J. and Belfort, M., (1988) Proc.Natl.Acad.Sci. (USA), 85 1151-1155.

Sjoberg, B.M., Hahne, S., Mathews, C.Z., Mathews, C., Rand, K.N. and Gait, M.J., (1986) EMBO J, 5 2031-2036.

Sogin, M.L. and Edman, J.C., (1989) Nucleic Acid Res., 17 5349-5359.

Sullivan, F.X. and Cech, T.R., (1985) Cell, 42 639-648.

Symons, R.H., (1992) Annu. Rev. Biochem., 61 641-671.

Symons, R.H., (1994) Curr. Opin. Struct. Biol., 4 322-330.

Tabak, H.F., Van der Horst, G., Osinga, K.A. and Asnberg, A.C., (1984) Cell, 39 623-629.

Tabak, H.F. and Grivell, L.A., (1986) Trends Genetics, 2 51-55.

Tabak, H.F., Van der Horst, G., Kamps, A.M.J.E. and Arnberg, A.C., (1987) Cell, 48 101-110.

Tanner, N.K. and Cech, T.R., (1987) Biochemistry, 26 3330-3340.

Tomich, P.K., Chiu, C.S., Wovcha, M.G. and Greenberg, G.R., (1974) J.Biol.chem., 249 7613-7622.

Tse, Y.C., Kirkegaard, K. and Wang, J.C., (1980) J.Biol.Chem., 255 5560-5565.

Tseng, M.-J., Hilfinger, J.M., Walsh, A. and Greenberg, G.R., (1988) J.Biol.Chem., 263 16242-16251.

Tseng, M.-J., Ping, H., Hilfinger, J.M. and Greenberg, G.R., (1990) J.Bacteriology, 172 6323-6332.

Tusch, T., Thomson, J.B. and Eckstein, F., (1995) Curr.Opin.Struct.Biol., 5 296-302.

Uhlenbeck, O.C., (1987) Nature, 328 596-600.

Van der Horst, G. and Tabak, H.F., (1985) Cell, 40 759-66.

Van der Veen, R., Arnberg, A.C., Van der Horst, G., Bonen, L., Tabak, H.F. and Grivell, L.A., (1986) Cell, 44 225-234.

Vinayak, R., Anderson, P., Mc Collum, C. and Hampel, A., (1992) Nucleic Acid Res., 20 1265-1269.

Wang, L.F. and Cech, T.R., (1992) Science, 256 526-529.

Wang, J.F., Downs, W.D. and Cech,T.R., (1993) Science, 260 504-508.

Waring, R.B., Davies, R.W., Scazzocchio, C. and Brown, T.A., (1982) Proc,Natl. Acad.Sci. (USA), 79 6332-6336.

Waring, R.B. and Davies, R.W., (1984) Gene, 28 277-291.

Waring, R.B., Towner, P., Minter, S.J. and Davies, R.W., (1986) Nature, 231 133-139.

Weiss-Brummer, B.,Holl, J., Schweyen, R.J., Rodel, G. and Kaudewitz, F., (1983) Cell, 33 195-202.

Wieringa, J.W., Hofer, E. and Weissmann, C., (1983) Cell, 37 915-923.

Williamson, C.L., Desai, N.M. and Burke, J.M., (1989) Nucleic Acid Res.,

17 675-89.

Wittop-Koning, T.H. and Schumperli, D., (1994) E.J.Biochem., 217 25-42.

Xu, M.Q. and Shub, D.A., (1989) Gene, 82 77-82.

Zaug, A.J., Grabowski, P.J. and Cech, T.R., (1983) Nature, 301 578-583.

Zaug, A.J., Kent, J.R. and Cech, T.R., (1984) Science, 224 574-578.

Zaug, A.J. and Cech,T.R., (1985) Science, 229 1060.

Zaug, A.J., Kent, J.R. and Cech, T.R., (1985) Biochemistry, 24 6211-6218.

Zaug, A.J. and Cech, T.R., (1986a) Science, 231 470.

Zaug, A.J. and Cech, T.R., (1986b) Biochemistry, 25 4478-4482.

LIST OF PUBLICATIONS AND PRESENTATIONS

1. **Khan, A.U.**, Lal, S.K. and Ahmad, M."Isolation and mapping of EMS-induced splicing defective point mutations within the intron of nrdB gene of bacteriophage T4". **Biochemical and Biophysical Research Communication,** (1998) 242 (1) 10-15.

2. **Khan, A.U.,** Lal, S.K. and Ahmad, M." Isolation and characterization of splicing defective mutations within the intron for the gene encoding the small subunit of T4 ribonucleotide reductase"." **FASEB J.** (1997) 11 (9) A-958.

3. **Khan, A.U.** Lal, S.K. and Ahmad, M. " Structural and functional studies of secondary structure domains of group I intron of nrdB gene of bacteriophage T4". **Invited lecture** in *National Symposium on Radiation and Molecular Biophysics* held at Mumbai during January 21-24, 1998.

4. **Khan, A.U.**, Lal, S.K. and Ahmad, M." Mapping of suppressor intron mutation to isolate extragenic revertants of T4 nrdB gene". **Poster presentation** in *66th Annual meeting of society of biological chemists (India)* held at Vishakapatnam during December 22-24, 1997.

5. **Khan, A.U.**, Lal, S.K. and Ahmad, M." Isolation and mapping of EMS-induced intron mutations of bacteriophage T4". **Poster presentation** in *65th Annual meeting of society of biological chemists (India)* held at Bangelore during November 20-23 1996.

SUMMARY

The *nrdB* gene of bacteriophage T4 codes for a small subunit of ribonucleotide reductase and contains 598 base-pair self splicing intron which is closely related to group I introns of T4 and eukaryotes. The strategy for isolation and mapping of EMS-induced splicing defective mutants within the intron of *nrdB* gene was based upon the appearance/disappearance of halo$^+$ phenotype on GPTG medium. Halo$^+$ or white halo phenotype implies is the excessive bacterial growth around the plaque due to the accumulation of uracil in *frd*1 mutants of T4 phage on infection to the OK305 strain of *E.coli* which is defective in its pyrimidine metabolism. Using this novel technique we have isolated a total of 432 EMS-induced *frd*1 double mutants (halo$^-$)which were further mapped within the *nrdB* gene by using marker rescue technique with the clone pJSS10 carrying the whole *nrdB* gene. A total of 156 mutants were mapped within the *nrdB* gene. The finer mapping of these 156 mutants was carried out by marker rescue technique with the pJBK1 clone and various subclones of *nrdB* intron. It was found that out of 31 intron mutants, 7 were mapped within the A region at the 5'end of intron, 11 were found to map within C region and remaining 13 were clustered near the 3'end of intron in the region D. None of the mutations was mapped in the region B which is having a non-conserved sequence.

The dot-blot splicing assay was performed to analyse splicing defective point mutations within the *nrdB* intron of bacteriophage T4. End-labelled DNA probes were designed such that their bindings were dependent on the splicing of the *nrdB* intron. Four DNA probes were used in this splicing assay. The wild type T4 RNA preparation was used as positive control to measure the splicing proficiency.

The present study led us to conclude the following:

(1) The conserved sequences in the *nrdB* intron like *td* and other group I introns were important in pre-mRNA splicing.

(2) The 5' and 3' domains of *nrdB* intron were more critical for splicing, especially the 3' end at which most of the splicing defective mutations were clustered.

(3) None of the mutations was found within the B region of *nrdB* intron which is outside the conserved P, Q, R, and S sequences.

(4) Dot-blot splicing assay suggested that all the mutations within the intron of *nrdB* gene were splicing defective and their defects were either at the 5'cutting site or at 3' cutting site, some of them have shown the defects at both the cutting sites. Interestingly, some of them were found to be only partially splicing defective.

(5) Exon sequences those immediately adjacent to the splice site were also important for splice site selection in splicing.

(6) ORF region of *nrdB* intron could also play an important role in the splicing of intron. The absence of the conserved region might imply that some sort of tertiary interactions between ORF and neighbouring conserved regions are involved to stabilize the RNA structure essential for splicing.

(7) It is clear from studies on *nrdB* intron that, it is a self splicing group I type intron with a core catalytic structure similar to other introns of its group.

In the second part of this study we isolated intragenic and extragenic revertants of our *frd*1*nrdB* double mutants of bacteriophage T4. These second site revertants were again obtained by EMS-induced mutagenesis. The basis of selection of these mutants was the inversion of the halo phenotype from halo$^-$ to halo$^+$ character. We have been able to isolate 23 revertants amongst them only one was found to be an extragenic. The isolation and mapping of these revertants were performed, similarly as was previously done in the case of splicing defective point mutations. The mapping results revealed that only 4 out of 22 intragenic reversions located within or near the intron of *nrdB* gene whereas remaining were probably either within the exon or near the ends of the exon I and exon II, i.e. the intron-exon junction. Further mapping of these four revertants within the intron was carried out by the usage of marker rescue technique with various sub-clones of intron. Out of 4 revertants, three were found to map near the primary mapping positions whereas one was mapped in the D region though, primarily it was mapped in the C region of intron.

The dot blot splicing assay was also performed with the same probes used as described earlier. It revealed that all the intragenic and extragenic revertants were found to be splicing defective except one intragenic revertant which was found to compensate the effect of its primary mutation whereas the isolated extragenic revertant could not restore splicing proficiency. These findings are suggestive of the following reasons:

(1) Most of the *nrdB* intragenic revertants were not mapped within the intron probably because, these mutants might be present on either extreme ends of intron or adjacent to the intron i.e. the end of exon I or start of exon II. Whereas the extragenic revertant that lies outside the *nrdB* gene would

probably disrupt the pyrimidine biosynthesis pathway resulting in the white halo phenotype.

(2) One intragenic suppressor mutant was detected as partially splicing proficient and hence this second site revertant would have been able to restore splicing activity of AM753 mutant to some extant.

(3) One revertant out of 23 was found to be extragenic. This second site mutation could not restore the splicing defect perhaps owing to a second mutation within the *frd* gene leading to accumulation of dUMP and thus producing weak halos by compensating for the effects of leaky *nrdB* mutation. This extragenic revertant did not affect splicing of *nrdB* RNA, but affected the halo phenotype.

(4) Mapping and dot-blot splicing assay suggested that the cleavage of the intron might involve a well defined secondary structure which on folding would probably attain a typical tertiary structure. These secondary and / or tertiary structures are crucial for the splicing of *nrdB* intron like other group I type introns.